

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:26:13 ; Search time 1092.5 Seconds
(without alignments)
88.945 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6

Sequence: 1 atgacg 6

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_yrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 6 | 100.0 | 19 | 9 | AA885444 am14h07.s |
| 2 | 6 | 100.0 | 19 | 13 | BM395830 5009-0-12 |
| 3 | 6 | 100.0 | 19 | 13 | BM396331 5009-0-12 |
| 4 | 6 | 100.0 | 19 | 13 | BM396331 5009-0-2 |
| 5 | 6 | 100.0 | 19 | 14 | BM399859 D44776 |
| 6 | 6 | 100.0 | 19 | 17 | D44776 HUMSUP214 |
| | | | | | AA317020 1M0035P18 |

| | | | | | | |
|----|---|-------|----|----|-----------|---------------------|
| 7 | 6 | 100.0 | 19 | 17 | A2457710 | A2457710 1M0261K07 |
| 8 | 6 | 100.0 | 19 | 17 | A2658087 | A2658087 1M0534J09 |
| 9 | 6 | 100.0 | 19 | 17 | A2663240 | A2663240 1M0542H18 |
| 10 | 6 | 100.0 | 19 | 17 | A2763729 | A2763729 1M0559N16 |
| 11 | 6 | 100.0 | 19 | 17 | A2861634 | A2861634 2M0168B18 |
| 12 | 6 | 100.0 | 19 | 17 | A2990851 | A2990851 2M0274E15 |
| 13 | 6 | 100.0 | 20 | 17 | A2500712 | A2500712 1M0339B13 |
| 14 | 6 | 100.0 | 20 | 17 | A2830285 | A2830285 2M0109N22 |
| 15 | 6 | 100.0 | 21 | 17 | A2455887 | A2455887 1M0258019 |
| 16 | 6 | 100.0 | 21 | 17 | A2662959 | A2662959 1M0542P18 |
| 17 | 6 | 100.0 | 21 | 17 | A2806895 | A2806895 2M0069E13 |
| 18 | 6 | 100.0 | 21 | 17 | A2949072 | A2949072 2M012D19 |
| 19 | 6 | 100.0 | 22 | 9 | A1001081 | A1001081 0S66A09.s |
| 20 | 6 | 100.0 | 22 | 17 | A2303795 | A2303795 1M0003I07 |
| 21 | 6 | 100.0 | 22 | 17 | A2314083 | A2314083 1M0030P11 |
| 22 | 6 | 100.0 | 22 | 17 | A2394940 | A2394940 1M0158P11 |
| 23 | 6 | 100.0 | 22 | 17 | A2651343 | A2651343 1M0522D06 |
| 24 | 6 | 100.0 | 22 | 17 | A2829109 | A2829109 2M0106L03 |
| 25 | 6 | 100.0 | 22 | 17 | BH863544 | BH863544 SALK_0940 |
| 26 | 6 | 100.0 | 22 | 17 | TA285H080 | TA285H080 T. brucei |
| 27 | 6 | 100.0 | 23 | 13 | BM395136 | BM395136 50072-2-7 |
| 28 | 6 | 100.0 | 23 | 13 | BM396053 | BM396053 5009-0-16 |
| 29 | 6 | 100.0 | 23 | 17 | A2307676 | A2307676 1M0009N24 |
| 30 | 6 | 100.0 | 23 | 17 | A2647964 | A2647964 1M0514F13 |
| 31 | 6 | 100.0 | 23 | 17 | A2808094 | A2808094 2M0071K24 |
| 32 | 6 | 100.0 | 23 | 17 | A2824304 | A2824304 2M0098N17 |
| 33 | 6 | 100.0 | 23 | 17 | A2838947 | A2838947 2M0134B21 |
| 34 | 6 | 100.0 | 23 | 17 | BH790464 | BH790464 SALK_0571 |
| 35 | 6 | 100.0 | 23 | 17 | BH790465 | BH790465 SALK_0571 |
| 36 | 6 | 100.0 | 23 | 17 | BH811030 | BH811030 SALK_0571 |
| 37 | 6 | 100.0 | 23 | 17 | BH811032 | BH811032 SALK_0571 |
| 38 | 6 | 100.0 | 24 | 17 | A2317925 | A2317925 1M0036S20 |
| 39 | 6 | 100.0 | 24 | 17 | A2510127 | A2510127 1M0354L21 |
| 40 | 6 | 100.0 | 24 | 17 | A2585771 | A2585771 1M0391L12 |
| 41 | 6 | 100.0 | 24 | 17 | A2624429 | A2624429 1M0463P10 |
| 42 | 6 | 100.0 | 24 | 17 | A2763574 | A2763574 1M0559P01 |
| 43 | 6 | 100.0 | 24 | 17 | A2784064 | A2784064 2M0026N20 |
| 44 | 6 | 100.0 | 24 | 17 | BH846515 | BH846515 SALK_0085 |
| 45 | 6 | 100.0 | 24 | 17 | TA75C020 | TA75C020 T. brucei |

ALIGNMENTS

RESULT 1
LOCUS AA885444 19 bp mRNA linear EST 04-JAN-1999
DEFINITION am14h07.s1 Soares_NFL.T.GBC.S1 Homo sapiens cDNA clone
IMAGE:1466845.3' similar to TR:000409 000409 CHECKPOINT SUPPRESSOR
1.1 mRNA sequence.

ACCESSION AA885444
VERSION AA885444.1 GI:2994521

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 19)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality

FEATURES
source
Possible reversed clone; similarity on wrong strand
Insert Length: 489 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1.19

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1466845"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT733-Pac (Pharmacia) with a modified polylinker; Site: 1; Not I; Site: 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CCAP GC91) were mixed and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 2 a 8 c 5 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 9; Length 19;
 Best Local Similarity 83.3%; Pred. No. 7.3e+05; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 12 ATGACT 7

RESULT 2
 BM395830 19 bp mRNA linear EST 17-JAN-2002
 LOCUS 5009-0-12-E03.t.1 Chilcoat/Turkewitz cDNA (large fraction)
 DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.
 ACCESSION BM395830
 VERSION BM395830.1 GI:18195863
 KEYWORDS EST.
 SOURCE Tetrahymena thermophila.
 ORGANISM Tetrahymena thermophila.
 Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida; Tetrahymenina; Tetrahymena.

REFERENCE 1 (bases 1 to 19)
 Turkewitz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel,J. and Klobutcher,L.
 EST from Tetrahymena thermophila, strain CU428.1, growing cells unpublished (2002)
 TITLE JOURNAL
 COMMENT Contact: Turkewitz AP
 Molecular Genetics and Cell Biology
 University of Chicago
 920 E. 58th Street, Chicago, IL 60637, USA
 Tel: 773 702 4374
 Fax: 773 702 3172
 Email: apturkew@midway.uchicago.edu
 Seq primer: T3.

FEATURES
 source
 1. 19
 Location/Qualifiers
 /organism="Tetrahymena thermophila"
 /strain="CU428.1"
 /db_xref="taxon:5911"
 /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
 /note="Vector: Bluescript2 SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT 5 a 5 c 6 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 13; Length 19;
 Best Local Similarity 83.3%; Pred. No. 7.3e+05; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 3 ATGACT 8

RESULT 3
 BM396331 19 bp mRNA linear EST 17-JAN-2002
 LOCUS 5009-0-2-E02.t.1 Chilcoat/Turkewitz cDNA (large fraction)
 DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.
 ACCESSION BM396331
 VERSION BM396331.1 GI:18196384
 KEYWORDS EST.
 SOURCE Tetrahymena thermophila.
 ORGANISM Tetrahymena thermophila.
 Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida; Tetrahymenina; Tetrahymena.

REFERENCE 1 (bases 1 to 19)
 Turkewitz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel,J. and Klobutcher,L.
 EST from Tetrahymena thermophila, strain CU428.1, growing cells unpublished (2002)
 TITLE JOURNAL
 COMMENT Contact: Turkewitz AP
 Molecular Genetics and Cell Biology
 University of Chicago
 920 E. 58th Street, Chicago, IL 60637, USA
 Tel: 773 702 4374
 Fax: 773 702 3172
 Email: apturkew@midway.uchicago.edu
 Seq primer: T3.

FEATURES
 source
 1. 19
 Location/Qualifiers
 /organism="Tetrahymena thermophila"
 /strain="CU428.1"
 /db_xref="taxon:5911"
 /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
 /note="Vector: Bluescript2 SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT 4 a 6 c 6 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 13; Length 19;
 Best Local Similarity 83.3%; Pred. No. 7.3e+05; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 2 ATGACT 7

RESULT 4
 BM399859 19 bp mRNA linear EST 17-JAN-2002
 LOCUS 5009-0-62-E04.t.2 Chilcoat/Turkewitz cDNA (large fraction)
 DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.
 ACCESSION BM399859
 VERSION BM399859.1 GI:18199912
 KEYWORDS EST.
 SOURCE Tetrahymena thermophila.
 ORGANISM Tetrahymena thermophila.
 Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida; Tetrahymenina; Tetrahymena.

REFERENCE 1 (bases 1 to 19)
 Turkewitz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel,J. and Klobutcher,L.
 EST from Tetrahymena thermophila, strain CU428.1, growing cells unpublished (2002)
 TITLE JOURNAL
 COMMENT Contact: Turkewitz AP
 Molecular Genetics and Cell Biology
 University of Chicago
 920 E. 58th Street, Chicago, IL 60637, USA
 Tel: 773 702 4374
 Fax: 773 702 3172
 Email: apturkew@midway.uchicago.edu
 Seq primer: T3.

```

FEATURES
  source
    1. 19
      Location/Qualifiers
        /organism="Tetrahymena thermophila"
        /strain="CU428.1"
        /db_xref="taxon:5911"
        /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
        /note="Vector: Bluescript2 SK+; Details on library
        preparation can be found in Chilcoat and Turkewitz (2001)
        Proc. Natl. Acad. Sci USA, 98: 8709-8713."
BASE COUNT      4 a      7 g      4 t
ORIGIN
  Query Match      100.0%; Score 6; DB 13; Length 19;
  Best Local Similarity 83.3%; Pred. No. 7.3e+05;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACU 6
    |||||:
    2 ATGACT 7
DB
  RESULT 5
  D44776      19 bp      mRNA      linear      EST 20-FEB-1998
  LOCUS      H0MSUPY214 Human brain cDNA Homo sapiens cDNA clone MFS1-S-2, mRNA
  DEFINITION sequence.
  ACCESSION  D44776
  VERSION    D44776.1 GI:1572251
  KEYWORDS  EST.
  SOURCE    human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  REFERENCE  1 (bases 1 to 19)
            Hadano,S., Ishida,Y., Tomiyasu,H., Yamamoto,K., Bates,G.P. and
            Ikeda,J.
            Transcript map of the human chromosome 4p16.3 consisting of 627
            DNA clones derived from 1 Mb of the Huntington's disease locus
            JOURNAL 97101646
            MEDLINE
            COMMENT Contact: Shinji Hadano
            Japan Science and Technology Corporation, Neurogenes Project, ICORP
            Univ. of Tokai School of Med.
            Bohseidai, Isehara, Kanagawa 259-1193, Japan
            Tel: 81-463-91-5095
            Fax: 81-463-91-4993
            Email: shinj@eng.med.u-tokai.ac.jp.
            Location/Qualifiers
              1. 19
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="MFS1-S-2"
                /clone_lib="Human brain cDNA"
                /tissue_type="brain"
                /note="Vector: pSPORT1; Site.1: M1ul; Site.2: NotI; mRNA
                was prepared from human testis of a 27 years old man. cDNA
                was prepared using a 15mer oligo dT anchored by two
                degenerated bases at its 3' end and containing a NotI site
                at its 5' end. The cDNA was cloned between SalI and NotI
                sites of pSPORT1. The M1ul-SalI fragment come from the
                adaptor used for the cloning. The 3' end is at the NotI
                site. cDNA corresponding to abundant species were
                eliminated from this library."
BASE COUNT      5 a      2 c      7 g      5 t
ORIGIN
  Query Match      100.0%; Score 6; DB 14; Length 19;
  Best Local Similarity 83.3%; Pred. No. 7.3e+05;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACU 6
    |||||:
    2 ATGACT 7
DB 13 ATGACT 18

```

```

  RESULT 6
  A2317020/c      19 bp      DNA      linear      GSS 29-SEP-2000
  LOCUS      IM0035P18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
  DEFINITION clone UUGC1M0035P18 F, DNA sequence.
  ACCESSION  A2317020
  VERSION    A2317020.1 GI:10365406
  KEYWORDS  GSS.
  SOURCE    house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
  REFERENCE  1 (bases 1 to 19)
            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            JOURNAL Unpublished (2000)
            COMMENT Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert length: 1000 Std Error: 0.00
            Plate: 0035 row: P column: 18
            Seq primer: CTTGTAAACGACGCCACT
            Class: plasmid ends
            High quality sequence stop: 19.
            Location/Qualifiers
              1. 19
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGC1M0035P18"
                /clone_lib="Mouse 10kb plasmid UUGC1M library"
                /sex="Male"
                /lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
                /note="Vector: pMD42nv; Purified genomic DNA from M.
                musculus C57BL/6J (male) was obtained from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a
                0.005 inch orifice at constant velocity. The sheared DNA
                was blunt end-repaired with T4 DNA polymerase and T4
                polynucleotide kinase. Adaptor oligonucleotides were
                ligated to the blunt ends in high molar excess. The
                adaptor DNA was purified and size-selected for a 9.5 to
                10.5 kb range using preparative agarose gel
                electrophoresis. Vector DNA was prepared from a derivative
                of pMD42 (91/4732114/9b)AF159072.1), a copy number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adaptor mouse DNA was annealed to
                adaptor vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
BASE COUNT      3 a      5 c      4 g      7 t
ORIGIN
  Query Match      100.0%; Score 6; DB 17; Length 19;
  Best Local Similarity 83.3%; Pred. No. 7.3e+05;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACU 6
    |||||:
    2 ATGACT 6
DB 11 ATGACT 6

```

RESULT 7
A2457710
LOCUS A2457710 19 bp DNA linear GSS 04-OCT-2000
DEFINITION IM0261K07F Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCGIM0261K07 F, DNA sequence.
ACCESSION A2457710
VERSION A2457710.1 GI:10615835
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0261 row: K column: 07
Seq primer: CCGTGTAAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES

source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCGIM0261K07"
/clone_1lb="Mouse 10kb plasmid UGCGIM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[bp]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 8 a 4 c 3 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 7.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
DB 5 ATGACT 10

RESULT 8
A2658087
LOCUS A2658087 19 bp DNA linear GSS 14-DEC-2000
DEFINITION IM0534J09R Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCGIM0534J09 R, DNA sequence.
ACCESSION A2658087
VERSION A2658087.1 GI:11795233
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0534 row: J column: 09
Seq primer: CACACAGAAACGATATGACC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES

source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCGIM0534J09"
/clone_1lb="Mouse 10kb plasmid UGCGIM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[bp]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 7 a 5 c 3 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 7.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
DB 2 ATGACT 7

RESULT 9
A2663240 19 bp DNA 1linear GSS 14-DEC-2000
LOCUS
DEFINITION
1M0542H18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0542H18 R, DNA sequence.
ACCESSION
A2663240
VERSION
A2663240.1 GI:11800386
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
Unpublished (2000)
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0542 row: H column: 18
Seq primer: CACACAGGAACACGCTATAC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0542H18"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-gold, T1-resistant, F-"
/note="Vector: PWD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
7 a 4 c 4 g 4 t
100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 7.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match
Best Local Similarity 83.3%; Pred. No. 7.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 10 ATGACT 15

RESULT 10
A2763729 19 bp DNA 1linear GSS 16-FEB-2001
LOCUS
DEFINITION
1M0559N16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0559N16 F, DNA sequence.
ACCESSION
A2763729
VERSION
A2763729.1 GI:12875056
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
Unpublished (2000)
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0559 row: N column: 16
Seq primer: CGTGTAAACGACGCGCAT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0559N16"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-gold, T1-resistant, F-"
/note="Vector: PWD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
7 a 3 c 3 g 6 t
100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 7.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match
Best Local Similarity 83.3%; Pred. No. 7.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 9 ATGACT 14

RESULT 11
 A2861634
 LOCUS 19 bp DNA linear GSS 21-FEB-2001
 DEFINITION 2M0168B18F Mouse 10kb plasmid UGCG1M library Mus musculus genomic
 accession A2861634
 clone U0CC2M0168B18 F, DNA sequence.
 VERSION A2861634.1 GI:13058150
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 19)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0168 row: B column: 18
 Seq primer: CGTGTAAACGACGCCACAG
 Class: plasmid ends
 High quality sequence stop: 19.

FEATURES

source
 1. 19
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UGCG2M0168B18"
 /clone_1lb="Mouse 10kb plasmid UGCG1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g114732114[gb]AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT

ORIGIN
 5 a 4 c 5 g 5 t
 Query Match 100.0%; Score 6; DB 17; Length 19;
 Best Local Similarity 83.3%; Pred. No. 7.3e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 8 ATGACT 13

RESULT 12
 A2990851
 LOCUS 19 bp DNA linear GSS 27-APR-2001
 DEFINITION 2M0274E15R Mouse 10kb plasmid UGCG2M library Mus musculus genomic
 accession A2990851
 clone U0CC2M0274E15 R, DNA sequence.
 VERSION A2990851
 clone U0CC2M0274E15 R, DNA sequence.
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 19)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0274 row: E column: 15
 Seq primer: CACACAGCAACAGCATGTGACC
 Class: plasmid ends
 High quality sequence stop: 19.

FEATURES

source
 1. 19
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UGCG2M0274E15"
 /clone_1lb="Mouse 10kb plasmid UGCG2M library"
 /sex="Female"
 /lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (female) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g114732114[gb]AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT

ORIGIN
 3 a 7 c 4 g 5 t
 Query Match 100.0%; Score 6; DB 17; Length 19;
 Best Local Similarity 83.3%; Pred. No. 7.3e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 7 ATGACT 12

RESULT 13
 A2500712/c 20 bp DNA linear GSS 05-OCT-2000
 LOCUS 1M0339B13F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
 DEFINITION clone UUGCIM0339B13 F, DNA sequence.
 ACCESSION A2500712
 VERSION A2500712.1 GI:106680801
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 20)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0339 row: B column: 13
 Seq primer: CGTTGTAAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers
 1..20
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCIM0339B13"
 /clone_1lb="Mouse 10kb plasmid UUGCIM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 7 a 6 c 1 g 6 t
 ORIGIN
 Query Match 100.0%; Score 6; DB 17; Length 20;
 Best Local Similarity 83.3%; Pred. No. 7.5e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 DB 11 ATGACT 6

RESULT 14
 A2830285 20 bp DNA linear GSS 20-FEB-2001
 LOCUS 2M0109N22F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
 DEFINITION clone UUGCIM0109N22 F, DNA sequence.
 ACCESSION A2830285
 VERSION A2830285.1 GI:13000193
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 20)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0109 row: N column: 22
 Seq primer: CGTTGTAAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers
 1..20
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCIM0109N22"
 /clone_1lb="Mouse 10kb plasmid UUGCIM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 6 c 6 g 4 t
 ORIGIN
 Query Match 100.0%; Score 6; DB 17; Length 20;
 Best Local Similarity 83.3%; Pred. No. 7.5e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 DB 1 ATGACT 6

RESULT 15
A2455887

LOCUS

DEFINITION 21 bp DNA linear GSS 04-OCT-2000
1M0258019F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

ACCESSION

A2455887
clone UUGC1M0258019 F, DNA sequence.

VERSION

A2455887.1 GI:10614012

KEYWORDS

GSS.
house mouse.

house mouse.

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.

AUTHORS

1 (bases 1 to 21)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily,
'M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0258 row: 0 column: 19
Seq primer: CGTGTAAACGACGCCAGTClass: plasmid ends
High quality sequence stop: 21.

FEATURES

Location/Qualifiers

1..21

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0258019"

/clone_1id="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (G114732114|gbl|AF129072.1) a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."BASE COUNT
ORIGIN

6 a 4 c 4 g 7 t

Query Match

100.0%; Score 6; DB 17; Length 21;

Best Local Similarity 83.3%; Pred. No. 7.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6

DB 3 ATGACT 8

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:56:27 ; Search time 333.5 Seconds

(without alignments)
523.589 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6
Sequence: 1 atgacu 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ov:*

8: gb_ov:*

9: gb_ov:*

10: gb_ov:*

11: gb_ov:*

12: gb_ov:*

13: gb_ov:*

14: gb_ov:*

15: gb_ov:*

16: gb_ov:*

17: gb_ov:*

18: gb_ov:*

19: gb_ov:*

20: gb_ov:*

21: gb_ov:*

22: gb_ov:*

23: gb_ov:*

24: gb_ov:*

25: gb_ov:*

26: gb_ov:*

27: gb_ov:*

28: gb_ov:*

29: gb_ov:*

30: gb_ov:*

31: gb_ov:*

32: gb_ov:*

33: gb_ov:*

34: gb_ov:*

35: gb_ov:*

36: gb_ov:*

37: gb_ov:*

38: gb_ov:*

39: gb_ov:*

40: gb_ov:*

41: gb_ov:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 6 | 100.0 | 8 | 6 | AX359003 | AX359003 Sequence |
| 2 | 6 | 100.0 | 10 | 6 | AR104790 | AR104790 Sequence |
| 3 | 6 | 100.0 | 10 | 6 | AX113032 | AX113032 Sequence |
| 4 | 6 | 100.0 | 10 | 6 | AX152230 | AX152230 Sequence |
| 5 | 6 | 100.0 | 10 | 6 | AX152244 | AX152244 Sequence |
| 6 | 6 | 100.0 | 10 | 6 | AX152408 | AX152408 Sequence |
| 7 | 6 | 100.0 | 10 | 6 | AX152409 | AX152409 Sequence |
| 8 | 6 | 100.0 | 10 | 6 | AX152854 | AX152854 Sequence |
| 9 | 6 | 100.0 | 10 | 6 | AX152855 | AX152855 Sequence |
| 10 | 6 | 100.0 | 10 | 6 | AX153355 | AX153355 Sequence |
| 11 | 6 | 100.0 | 10 | 6 | AX239915 | AX239915 Sequence |
| 12 | 6 | 100.0 | 10 | 6 | AX301406 | AX301406 Sequence |
| 13 | 6 | 100.0 | 10 | 6 | AX301520 | AX301520 Sequence |
| 14 | 6 | 100.0 | 10 | 6 | AX301566 | AX301566 Sequence |
| 15 | 6 | 100.0 | 10 | 6 | AX301578 | AX301578 Sequence |
| 16 | 6 | 100.0 | 10 | 6 | AX377149 | AX377149 Sequence |
| 17 | 6 | 100.0 | 11 | 6 | AX022911 | AX022911 Sequence |
| 18 | 6 | 100.0 | 11 | 6 | AX022930 | AX022930 Sequence |
| 19 | 6 | 100.0 | 11 | 6 | AX022949 | AX022949 Sequence |
| 20 | 6 | 100.0 | 11 | 6 | AX030499 | AX030499 Sequence |
| 21 | 6 | 100.0 | 11 | 6 | AX030518 | AX030518 Sequence |
| 22 | 6 | 100.0 | 11 | 6 | AX030537 | AX030537 Sequence |
| 23 | 6 | 100.0 | 11 | 6 | AX471026 | AX471026 Sequence |
| 24 | 6 | 100.0 | 11 | 6 | AX471139 | AX471139 Sequence |
| 25 | 6 | 100.0 | 11 | 6 | AX471204 | AX471204 Sequence |
| 26 | 6 | 100.0 | 11 | 6 | AX471618 | AX471618 Sequence |
| 27 | 6 | 100.0 | 11 | 6 | AX471768 | AX471768 Sequence |
| 28 | 6 | 100.0 | 11 | 6 | AX472198 | AX472198 Sequence |
| 29 | 6 | 100.0 | 11 | 6 | I28557 | I28557 Sequence 10 |
| 30 | 6 | 100.0 | 11 | 6 | I56092 | I56092 Sequence 3 |
| 31 | 6 | 100.0 | 11 | 6 | I58719 | I58719 Sequence 10 |
| 32 | 6 | 100.0 | 12 | 6 | AR058577 | AR058577 Sequence |
| 33 | 6 | 100.0 | 12 | 6 | AR058578 | AR058578 Sequence |
| 34 | 6 | 100.0 | 12 | 6 | AR167775 | AR167775 Sequence |
| 35 | 6 | 100.0 | 12 | 6 | AR193085 | AR193085 Sequence |
| 36 | 6 | 100.0 | 12 | 6 | AX464720 | AX464720 Sequence |
| 37 | 6 | 100.0 | 12 | 6 | E29659 | E29659 Method for |
| 38 | 6 | 100.0 | 12 | 6 | E38765 | E38765 Method and |
| 39 | 6 | 100.0 | 12 | 6 | E64191 | E64191 Method for |
| 40 | 6 | 100.0 | 13 | 6 | I04302 | I04302 Sequence 15 |
| 41 | 6 | 100.0 | 13 | 6 | AR069833 | AR069833 Sequence |
| 42 | 6 | 100.0 | 13 | 6 | AR149120 | AR149120 Sequence |
| 43 | 6 | 100.0 | 13 | 6 | AX048318 | AX048318 Sequence |
| 44 | 6 | 100.0 | 13 | 6 | AX266958 | AX266958 Sequence |
| 45 | 6 | 100.0 | 13 | 6 | AX300869 | AX300869 Sequence |

ALIGNMENTS

RESULT 1

AX359003

LOCUS AX359003

DEFINITION Sequence 10 from Patent WO0183737.

ACCESSION AX359003

VERSION AX359003.1 GI:18675402

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1

Nagy,E., Tuboly,T. and Nagy,M.

Porcine adenovirus vaccine

Patent: WO 0183737-A 10 08-NOV-2001;

UNIVERSITY OF GUELPH (CA)

8 bp

DNA

Linear

PAT 13-FEB-2002

```
FEATURES
  source
    1. .8
    /organism="synthetic construct"
    /db_xref="taxon:32630"
    /note="Fragment"
BASE COUNT
  3 a 1 c 2 g 2 t
ORIGIN

Query Match
  100.0%; Score 6; DB 6; Length 8;
Best Local Similarity 83.3%; Pred. No. 3.6e+09;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
  |||||
  1 ATGACT 6
DB 1 ATGACT 6

RESULT 2
LOCUS ARI04790 10 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 87 from patent US 6093811.
ACCESSION ARI04790
VERSION ARI04790.1 GI:12817498
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Bennett,C.Frank, and Mirabelli,C.K.
TITLE Oligonucleotide modulation of cell adhesion
JOURNAL Patent: US 6093811-A 87 25-JUL-2000;
FEATURES
  source
    1. .10
    /organism="unknown"
BASE COUNT
  4 a 2 c 1 g 3 t
ORIGIN

Query Match
  100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
  |||||
  1 ATGACT 10
DB 5 ATGACT 10

RESULT 3
LOCUS AXI13032 10 bp DNA linear PAT 01-MAY-2001
DEFINITION Sequence 79 from Patent WO0127267.
ACCESSION AXI13032
VERSION AXI13032.1 GI:13939467
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE 1 (bases 1 to 10)
AUTHORS Adams,E., Waldmann,H., Cobbold,S. and Zelenitska,D.
TITLE Genes differentially expressed in T1 cells and their use in the
JOURNAL manufacture of immunoregulatory compositions
JOURNAL Patent: WO 0127267-A 79 19-APR-2001;
FEATURES
  source
    1. .10
    /organism="Mus sp."
    /db_xref="taxon:10095"
BASE COUNT
  4 a 1 c 2 g 3 t
ORIGIN

Query Match
  100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
  |||||
  1 ATGACT 5
DB 10 ATGACT 5

RESULT 4
LOCUS AXI52230/c 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 145 from Patent WO0138577.
ACCESSION AXI52230
VERSION AXI52230.1 GI:14533881
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 145 31-MAY-2001;
FEATURES
  source
    1. .10
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT
  3 a 2 c 2 g 3 t
ORIGIN

Query Match
  100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
  |||||
  1 ATGACT 4
DB 9 ATGACT 4

RESULT 5
LOCUS AXI52244/c 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 159 from Patent WO0138577.
ACCESSION AXI52244
VERSION AXI52244.1 GI:14533895
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 159 31-MAY-2001;
FEATURES
  source
    1. .10
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT
  3 a 2 c 2 g 3 t
ORIGIN

Query Match
  100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
  |||||
  1 ATGACT 5
DB 10 ATGACT 5

RESULT 6
LOCUS AXI52408 10 bp DNA linear PAT 22-JUN-2001
```

DEFINITION Sequence 323 from Patent WO0138577.
ACCESSION AX152408
VERSION AX152408.1 GI:14534059
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 323 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 1 ATGACT 6

RESULT 7
AX152409 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS
DEFINITION Sequence 324 from Patent WO0138577.
ACCESSION AX152409
VERSION AX152409.1 GI:14534060
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 324 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 1 ATGACT 6

RESULT 8
AX152854/c 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS
DEFINITION Sequence 769 from Patent WO0138577.
ACCESSION AX152854
VERSION AX152854.1 GI:14534505
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)

AUTHORS Velculescu, V.E., Vogelstein, B. and Kinzler, K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 769 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 7 ATGACT 2

RESULT 9
AX152855/c 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS
DEFINITION Sequence 770 from Patent WO0138577.
ACCESSION AX152855
VERSION AX152855.1 GI:14534506
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 770 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 7 ATGACT 2

RESULT 10
AX153355/c 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS
DEFINITION Sequence 1270 from Patent WO0138577.
ACCESSION AX153355
VERSION AX153355.1 GI:14535006
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 1270 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2 a 1 c 4 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 83.3%; Pred. No. 3.2e+06;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 8 ATGACT 3

RESULT 11
 AX239915 10 bp DNA linear PAT 26-SEP-2001
 LOCUS Sequence 42 from Patent WO0164958.
 DEFINITION AX239915
 ACCESSION AX239915
 VERSION AX239915.1 GI:15797517

KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 synthetic construct.
 artificial sequences.

REFERENCE
 AUTHORS Dempsey, R.O., Gall, A.A., Lokhov, S.G., Afonina, I.A., Singer, M.J.,
 Kutayvin, I.V. and Vermeulen, N.M.
 Modified oligonucleotides for mismatch discrimination
 Patent: WO 0164958-A 42 07-SEP-2001;
 Epoch Biosciences, Inc. (US)

TITLE
 JOURNAL
 location/Qualifiers

FEATURES
 source
 1..10
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="duplex complement 7"

BASE COUNT
 ORIGIN
 4 a 2 c 2 g 2 t

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 83.3%; Pred. No. 3.2e+06;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 5 ATGACT 10

RESULT 12
 AX301406/c 10 bp DNA linear PAT 30-NOV-2001
 LOCUS Sequence 120 from Patent WO0185941.
 DEFINITION AX301406
 ACCESSION AX301406
 VERSION AX301406.1 GI:17382489

KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Versteeg, R. and Caron, H.N.
 MYC targets
 Patent: WO 0185941-A 120 15-NOV-2001;
 Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)

TITLE
 JOURNAL
 location/Qualifiers

FEATURES
 source
 1..10
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT
 ORIGIN
 4 a 1 c 2 g 3 t

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 83.3%; Pred. No. 3.2e+06;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 1 ATGACT 3

Db 8 ATGACT 3

RESULT 13
 AX301520 10 bp DNA linear PAT 30-NOV-2001
 LOCUS Sequence 234 from Patent WO0185941.
 DEFINITION AX301520
 ACCESSION AX301520
 VERSION AX301520.1 GI:17382603

KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Versteeg, R. and Caron, H.N.
 MYC targets
 Patent: WO 0185941-A 234 15-NOV-2001;
 Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)

TITLE
 JOURNAL
 location/Qualifiers

FEATURES
 source
 1..10
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT
 ORIGIN
 4 a 1 c 1 g 4 t

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 83.3%; Pred. No. 3.2e+06;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 2 ATGACT 7

RESULT 14
 AX301566 10 bp DNA linear PAT 30-NOV-2001
 LOCUS Sequence 280 from Patent WO0185941.
 DEFINITION AX301566
 ACCESSION AX301566
 VERSION AX301566.1 GI:17382649

KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Versteeg, R. and Caron, H.N.
 MYC targets
 Patent: WO 0185941-A 280 15-NOV-2001;
 Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)

TITLE
 JOURNAL
 location/Qualifiers

FEATURES
 source
 1..10
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT
 ORIGIN
 4 a 1 c 1 g 4 t

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 83.3%; Pred. No. 3.2e+06;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 2 ATGACT 7

RESULT 15
 AX301578/c 10 bp DNA linear PAT 30-NOV-2001
 LOCUS Sequence 292 from Patent WO0185941.
 DEFINITION AX301578
 ACCESSION AX301578
 VERSION AX301578.1 GI:17382661

KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Versteeg, R. and Caron, H.N.
TITLE Myc targets
JOURNAL Patent: WO 0185941-A 292 15-NOV-2001;
Academisch ziekenhuis bij de Universiteit van Amsterdam (NL)
Location/Qualifiers

FEATURES
source
1. .10
/db_xref="taxon:9606"
/db_xref="taxon:9606"

BASE COUNT 4 a 1 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||:
Db 8 ATGACT 3

Search completed: May 21, 2003, 05:12:44
Job time : 339 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:52:32 ; Search time 156 Seconds
(without alignments)
86.615 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6
Sequence: 1 atgacu 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.GeneSeq_101002.*
1: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 100.0 | 8 | 21 | AAA80753 | A. thaliana primer |
| 2 | 100.0 | 8 | 21 | AAA80845 | A. thaliana primer |
| 3 | 100.0 | 24 | 16 | AAK87314 | GCN4 yeast recogni |
| 4 | 100.0 | 10 | 16 | AAO96636 | HIV-1 NL4-3 nef ge |
| 5 | 100.0 | 10 | 16 | AAO96637 | HIV-1 NL4-3 nef ge |
| 6 | 100.0 | 10 | 16 | AAO96638 | HIV-1 NL4-3 nef ge |
| 7 | 100.0 | 10 | 16 | AAO96639 | HIV-1 NL4-3 nef ge |
| 8 | 100.0 | 10 | 16 | AAO96640 | HIV-1 NL4-3 nef ge |
| 9 | 100.0 | 10 | 19 | AAV50231 | Yeast tag for addi |

| 10 | 6 | 100.0 | 10 | 20 | AAI18647 | p53 serial analysi |
|----|--|---|--|--|---------------------------------------|---------------------|
| 11 | 6 <td>100.0 <td>10 <td>20 <td>AAI18610 <td>p53 serial analysi</td> </td></td></td></td> | 100.0 <td>10 <td>20 <td>AAI18610 <td>p53 serial analysi</td> </td></td></td> | 10 <td>20 <td>AAI18610 <td>p53 serial analysi</td> </td></td> | 20 <td>AAI18610 <td>p53 serial analysi</td> </td> | AAI18610 <td>p53 serial analysi</td> | p53 serial analysi |
| 12 | 6 <td>100.0 <td>10 <td>21 <td>AAA76009 <td>DNA sequence analys</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAA76009 <td>DNA sequence analys</td> </td></td></td> | 10 <td>21 <td>AAA76009 <td>DNA sequence analys</td> </td></td> | 21 <td>AAA76009 <td>DNA sequence analys</td> </td> | AAA76009 <td>DNA sequence analys</td> | DNA sequence analys |
| 13 | 6 <td>100.0 <td>10 <td>21 <td>AAA76025 <td>DNA sequence analys</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAA76025 <td>DNA sequence analys</td> </td></td></td> | 10 <td>21 <td>AAA76025 <td>DNA sequence analys</td> </td></td> | 21 <td>AAA76025 <td>DNA sequence analys</td> </td> | AAA76025 <td>DNA sequence analys</td> | DNA sequence analys |
| 14 | 6 <td>100.0 <td>10 <td>21 <td>AAZ77646 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ77646 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ77646 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ77646 <td>Human dendritic ce</td> </td> | AAZ77646 <td>Human dendritic ce</td> | Human dendritic ce |
| 15 | 6 <td>100.0 <td>10 <td>21 <td>AAZ77660 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ77660 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ77660 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ77660 <td>Human dendritic ce</td> </td> | AAZ77660 <td>Human dendritic ce</td> | Human dendritic ce |
| 16 | 6 <td>100.0 <td>10 <td>21 <td>AAZ77670 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ77670 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ77670 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ77670 <td>Human dendritic ce</td> </td> | AAZ77670 <td>Human dendritic ce</td> | Human dendritic ce |
| 17 | 6 <td>100.0 <td>10 <td>21 <td>AAZ78367 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ78367 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ78367 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ78367 <td>Human dendritic ce</td> </td> | AAZ78367 <td>Human dendritic ce</td> | Human dendritic ce |
| 18 | 6 <td>100.0 <td>10 <td>21 <td>AAZ78828 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ78828 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ78828 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ78828 <td>Human dendritic ce</td> </td> | AAZ78828 <td>Human dendritic ce</td> | Human dendritic ce |
| 19 | 6 <td>100.0 <td>10 <td>21 <td>AAZ79615 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ79615 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ79615 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ79615 <td>Human dendritic ce</td> </td> | AAZ79615 <td>Human dendritic ce</td> | Human dendritic ce |
| 20 | 6 <td>100.0 <td>10 <td>21 <td>AAZ79868 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ79868 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ79868 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ79868 <td>Human dendritic ce</td> </td> | AAZ79868 <td>Human dendritic ce</td> | Human dendritic ce |
| 21 | 6 <td>100.0 <td>10 <td>21 <td>AAZ81560 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ81560 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ81560 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ81560 <td>Human dendritic ce</td> </td> | AAZ81560 <td>Human dendritic ce</td> | Human dendritic ce |
| 22 | 6 <td>100.0 <td>10 <td>21 <td>AAZ82255 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ82255 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ82255 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ82255 <td>Human dendritic ce</td> </td> | AAZ82255 <td>Human dendritic ce</td> | Human dendritic ce |
| 23 | 6 <td>100.0 <td>10 <td>21 <td>AAZ83058 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ83058 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ83058 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ83058 <td>Human dendritic ce</td> </td> | AAZ83058 <td>Human dendritic ce</td> | Human dendritic ce |
| 24 | 6 <td>100.0 <td>10 <td>21 <td>AAZ83146 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ83146 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ83146 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ83146 <td>Human dendritic ce</td> </td> | AAZ83146 <td>Human dendritic ce</td> | Human dendritic ce |
| 25 | 6 <td>100.0 <td>10 <td>21 <td>AAZ83365 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ83365 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ83365 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ83365 <td>Human dendritic ce</td> </td> | AAZ83365 <td>Human dendritic ce</td> | Human dendritic ce |
| 26 | 6 <td>100.0 <td>10 <td>21 <td>AAZ84650 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ84650 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ84650 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ84650 <td>Human dendritic ce</td> </td> | AAZ84650 <td>Human dendritic ce</td> | Human dendritic ce |
| 27 | 6 <td>100.0 <td>10 <td>21 <td>AAZ85998 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ85998 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ85998 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ85998 <td>Human dendritic ce</td> </td> | AAZ85998 <td>Human dendritic ce</td> | Human dendritic ce |
| 28 | 6 <td>100.0 <td>10 <td>21 <td>AAZ86484 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>21 <td>AAZ86484 <td>Human dendritic ce</td> </td></td></td> | 10 <td>21 <td>AAZ86484 <td>Human dendritic ce</td> </td></td> | 21 <td>AAZ86484 <td>Human dendritic ce</td> </td> | AAZ86484 <td>Human dendritic ce</td> | Human dendritic ce |
| 29 | 6 <td>100.0 <td>10 <td>22 <td>AAZ12883 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAZ12883 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAZ12883 <td>Human dendritic ce</td> </td></td> | 22 <td>AAZ12883 <td>Human dendritic ce</td> </td> | AAZ12883 <td>Human dendritic ce</td> | Human dendritic ce |
| 30 | 6 <td>100.0 <td>10 <td>22 <td>AAH63505 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH63505 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH63505 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH63505 <td>Human dendritic ce</td> </td> | AAH63505 <td>Human dendritic ce</td> | Human dendritic ce |
| 31 | 6 <td>100.0 <td>10 <td>22 <td>AAH63519 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH63519 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH63519 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH63519 <td>Human dendritic ce</td> </td> | AAH63519 <td>Human dendritic ce</td> | Human dendritic ce |
| 32 | 6 <td>100.0 <td>10 <td>22 <td>AAH63483 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH63483 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH63483 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH63483 <td>Human dendritic ce</td> </td> | AAH63483 <td>Human dendritic ce</td> | Human dendritic ce |
| 33 | 6 <td>100.0 <td>10 <td>22 <td>AAH63484 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH63484 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH63484 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH63484 <td>Human dendritic ce</td> </td> | AAH63484 <td>Human dendritic ce</td> | Human dendritic ce |
| 34 | 6 <td>100.0 <td>10 <td>22 <td>AAH63929 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH63929 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH63929 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH63929 <td>Human dendritic ce</td> </td> | AAH63929 <td>Human dendritic ce</td> | Human dendritic ce |
| 35 | 6 <td>100.0 <td>10 <td>22 <td>AAH63930 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH63930 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH63930 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH63930 <td>Human dendritic ce</td> </td> | AAH63930 <td>Human dendritic ce</td> | Human dendritic ce |
| 36 | 6 <td>100.0 <td>10 <td>22 <td>AAH64430 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH64430 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH64430 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH64430 <td>Human dendritic ce</td> </td> | AAH64430 <td>Human dendritic ce</td> | Human dendritic ce |
| 37 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 38 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 39 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 40 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 41 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 42 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 43 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 44 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |
| 45 | 6 <td>100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td></td> | 100.0 <td>10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td></td> | 10 <td>22 <td>AAH20008 <td>Human dendritic ce</td> </td></td> | 22 <td>AAH20008 <td>Human dendritic ce</td> </td> | AAH20008 <td>Human dendritic ce</td> | Human dendritic ce |

ALIGNMENTS

RESULT 1
AAA80753/C
ID AAA80753 standard; DNA: 8 BP.
XX
AC AAA80753;
XX
24-NOV-2000 (first entry)
XX
DE A. thaliana primer walking octamer SEQ ID NO: 66.
XX
XX Primer walking; octamer; primer; DNA sequencing; PCR; ss.
OS Arabidopsis thaliana.
XX
PN US6083695-A.
XX
PD 04-JUL-2000.
XX
PF 21-MAY-1997; 97US-0859954.
XX
PR 15-APR-1996; 96US-0632782.
XX
XX (UYHO-) UNIV HOUSTON.
PA (HARD/) HARDIN S H.
XX
PI Hardin PE, Hardin SH, Homayouni R;
XX
XX WPI: 2000-474852/41.
XX
PT Sequencing an unknown DNA molecule for the polymerase chain reaction
PT and other primer processes comprises primer walking of octamer
PT oligonucleotides

XX Example 8; Column 57-58; 161pp; English.
PS
XX
CC This invention describes a novel method for sequencing an unknown DNA
CC molecule which comprises selecting a library primer from an octamer
CC oligonucleotide library consisting of 48 8-bp sequences and
CC corresponding complementary sequences, where the library primer is
CC complementary to a known sequence adjacent to the unknown sequence or
CC is complementary to a sequence in a known extension product. The method
CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
CC which make use of primers. The octamers are used to identify coding
CC sequences. Primer walking using the octamer libraries is advantageous
CC over other sequencing methods because it does not require multiple
CC cloning steps nor subsequent template preparations, and it is a
CC directed and methodical approach. AAA80688-A81253 represent the octamer
CC primers used in the primer walking method of the invention.
XX
SQ Sequence 8 BP; 2 A; 1 C; 3 G; 2 T; 0 other;
Query Match 100.0%; Score 6; DB 21; Length 8;
Best Local Similarity 83.3%; Pred. No. 2.7e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACU 6
DB 8 ATGACT 3
RESULT 2
AAA80845
ID AAA80845 standard; DNA; 8 BP.
XX
XX AAA80845;
AC
XX
XX 24-NOV-2000 (first entry)
DT
XX
XX A. thaliana primer walking octamer SEQ ID NO: 158.
DE
XX
XX Primer walking; octamer; primer; DNA sequencing; PCR; ss.
KM
XX
XX Arabidopsis thaliana.
OS
XX
XX US6083695-A.
PN
XX
XX 04-JUL-2000.
PD
XX
XX 21-MAY-1997; 97US-0859954.
PF
XX
XX 15-APR-1996; 96US-0632782.
PR
XX
XX (UYHO-) UNIV HOUSTON.
PA (HARD/) HARDIN S H.
PI Hardin PE, Hardin SH, Homayouni R;
XX
XX WPI; 2000-474852/41.
DR
XX
XX Sequencing an unknown DNA molecule for the polymerase chain reaction
PT and other primer processes comprises primer walking of octamer
PT oligonucleotides -
XX
XX Example 8; Column 105-106; 161pp; English.
PS
XX
XX This invention describes a novel method for sequencing an unknown DNA
CC molecule which comprises selecting a library primer from an octamer
CC oligonucleotide library consisting of 48 8-bp sequences and
CC corresponding complementary sequences, where the library primer is
CC complementary to a known sequence adjacent to the unknown sequence or
CC is complementary to a sequence in a known extension product. The method
CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
CC which make use of primers. The octamers are used to identify coding
CC sequences. Primer walking using the octamer libraries is advantageous
CC over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
CC directed and methodical approach. AAA80688-A81253 represent the octamer
CC primers used in the primer walking method of the invention.
XX
SQ Sequence 8 BP; 2 A; 2 C; 2 G; 2 T; 0 other;
Query Match 100.0%; Score 6; DB 21; Length 8;
Best Local Similarity 83.3%; Pred. No. 2.7e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACU 6
DB 2 ATGACT 7
RESULT 3
ABK87314
ID ABK87314 standard; DNA; 9 BP.
XX
XX ABK87314;
AC
XX
XX 24-SEP-2002 (first entry)
DT
XX
XX GCN4 yeast recognition sequence #1.
DE
XX
XX Nucleic acid detection; GCN4; lambda repressor; ss; yeast.
KM
XX
XX Saccharomyces.
OS
XX
XX WO200244326-A2.
PN
XX
XX 06-JUN-2002.
PD
XX
XX 26-NOV-2001; 2001MO-US44215.
PF
XX
XX 30-NOV-2000; 2000US-0728574.
PR
XX
XX (STRA-) STRATAGENE.
PA
XX
XX Sorge JA, Whalen AM;
PI
XX
XX WPI; 2002-508503/54.
DR
XX
XX Detecting/measuring target nucleic acid, by forming cleavage structure
PT by incubating target nucleic acid with probe having binding moiety,
PT cleaving structure to release nucleic acid and detecting released
PT fragments -
XX
XX Disclosure; Page 75; 157pp; English.
PS
XX
XX This invention relates to a novel method for detecting/measuring a
CC target nucleic acid. The method comprises forming a cleavage structure
CC by incubating the target sequence with a probe comprising a binding
CC moiety and a secondary structure that changes upon binding of the probe
CC to the target, cleaving the cleavage structure to release a nucleic
CC acid fragment, and detecting and/or measuring the fragment captured by
CC binding of the binding moiety to a capture element on a solid support.
CC The method of the invention is useful for detecting or measuring a
CC target nucleic acid and are useful for generating a signal indicative of
CC the presence of the target nucleic acid in a sample. Another method of
CC the invention is useful for simultaneously forming a cleavage structure,
CC amplifying the target nucleic acid in a sample and cleaving the cleavage
CC structure. The method does not require multiple steps, subsequent
CC amplification process, and allows for concurrent amplification and
CC detection of target nucleic acid in a sample. The present sequence
CC represents the yeast GCN4 recognition sequence shown in the
CC specification.
XX
SQ Sequence 9 BP; 3 A; 2 C; 1 G; 3 T; 0 other;
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 OS |||||:
 PN 1 ATGACT 6
 DB

RESULT 4
 AAO96636 standard; DNA; 10 BP.

XX AAO96636;

XX 22-MAR-1996 (first entry)

XX HIV-1 NL4-3 nef gene nucleotide deletion 231.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX Human Immunodeficiency virus type 1.

XX WO9521912-A1.

XX 17-AUG-1995.

XX 14-FEB-1995; 95WO-AU00063.

XX 23-DEC-1994; 94AU-0000284.

XX 14-FEB-1994; 94AU-0003864.

XX 21-FEB-1994; 94AU-0004002.

XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

XX or LTR region - can be used in a vaccine to inhibit/reduce

XX productive infection in an individual by a pathogenic strain

XX Claim 13; Page 191; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

XX or more deannucleotides (AAO96406-097018) from the nef gene and/or

XX 1 or more deannucleotides (AAO97019-097166) from the LTR region; the

XX sequence of AAO96406 corresponds to nucleotides 1-10 of the nef gene

XX (AAO96141). The resulting avirulent HIV strains are still capable of

XX inducing an immune response in humans, and enable the generation of

XX therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 4 A; 3 C; 1 G; 2 T; 0 other;

XX Query Match 100.0%; Score 6; DB 16; Length 10;

XX Best Local Similarity 83.3%; Pred. No. 1.6e+05;

XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 ATGACU 6

XX |||||:
 XX 5 ATGACT 10
 DB

RESULT 5
 AAO96637 standard; DNA; 10 BP.

XX AAO96637;

XX 22-MAR-1996 (first entry)

XX HIV-1 NL4-3 nef gene nucleotide deletion 232.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX Human Immunodeficiency virus type 1.
 XX WO9521912-A1.
 XX 17-AUG-1995.
 XX 14-FEB-1995; 95WO-AU00063.
 XX 23-DEC-1994; 94AU-0000284.
 XX 14-FEB-1994; 94AU-0003864.
 XX 21-FEB-1994; 94AU-0004002.

XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 XX or LTR region - can be used in a vaccine to inhibit/reduce
 XX productive infection in an individual by a pathogenic strain

XX Claim 13; Page 191; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

XX or more deannucleotides (AAO96406-097018) from the nef gene and/or

XX 1 or more deannucleotides (AAO97019-097166) from the LTR region; the

XX sequence of AAO96406 corresponds to nucleotides 1-10 of the nef gene

XX (AAO96141). The resulting avirulent HIV strains are still capable of

XX inducing an immune response in humans, and enable the generation of

XX therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 3 A; 3 C; 1 G; 3 T; 0 other;

XX Query Match 100.0%; Score 6; DB 16; Length 10;

XX Best Local Similarity 83.3%; Pred. No. 1.6e+05;

XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 ATGACU 6

XX |||||:
 XX 4 ATGACT 9
 DB

RESULT 6
 AAO96638 standard; DNA; 10 BP.

XX AAO96638;

XX 22-MAR-1996 (first entry)

XX HIV-1 NL4-3 nef gene nucleotide deletion 233.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX Human Immunodeficiency virus type 1.

XX WO9521912-A1.

XX 17-AUG-1995.

XX 14-FEB-1995; 95WO-AU00063.

XX 23-DEC-1994; 94AU-0000284.

XX 14-FEB-1994; 94AU-0003864.

XX 21-FEB-1994; 94AU-0004002.

XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX WPI; 1995-293115/38.

XX WPI; 1995-293115/38.
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX
 PS Claim 13; Page 191; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decaunucleotides (AAQ96406-Q97018) from the nef gene and/or
 CC 1 or more decaunucleotides (AAQ97019-Q97166) from the LTR region; the
 CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
 CC (AAQ96141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 2 C; 1 G; 3 T; 0 other;
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.6e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACU 6
 Db 3 ATGACT 8
 RESULT 7
 ID AAQ96639 standard; DNA; 10 BP.
 XX
 AC AAQ96639;
 XX
 DT 22-MAR-1996 (first entry)
 XX
 DE HIV-1 NL4-3 nef gene nucleotide deletion 234.
 XX
 KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PF 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.
 XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 PI WPI; 1995-293115/38.
 DR
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX
 PS Claim 13; Page 191; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decaunucleotides (AAQ96406-Q97018) from the nef gene and/or
 CC 1 or more decaunucleotides (AAQ97019-Q97166) from the LTR region; the
 CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
 CC (AAQ96141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX

SQ Sequence 10 BP; 4 A; 2 C; 1 G; 3 T; 0 other;
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.6e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACU 6
 Db 2 ATGACT 7
 RESULT 8
 ID AAQ96640 standard; DNA; 10 BP.
 XX
 AC AAQ96640;
 XX
 DT 22-MAR-1996 (first entry)
 XX
 DE HIV-1 NL4-3 nef gene nucleotide deletion 235.
 XX
 KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PF 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.
 XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 PI WPI; 1995-293115/38.
 DR
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX
 PS Claim 13; Page 191; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decaunucleotides (AAQ96406-Q97018) from the nef gene and/or
 CC 1 or more decaunucleotides (AAQ97019-Q97166) from the LTR region; the
 CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
 CC (AAQ96141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 2 C; 1 G; 3 T; 0 other;
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.6e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACU 6
 Db 1 ATGACT 6
 RESULT 9
 ID AAV50231 standard; DNA; 10 BP.
 XX
 AC AAV50231;
 XX

```

DE DT 21-Oct-1998 (first entry)
XX
XX Yeast tag for additional NORF chromosome 15 tag position 32081.
XX
XX Yeast; Saccharomyces cerevisiae; transcriptome; cell cycle;
KW regulation; eukaryotic cell; antifungal; SAGE tag; gene expression;
KW serial analysis of gene expression; probe; ss.
XX
XX Saccharomyces cerevisiae.
OS Synthetic.
PN
XX MO9832847-A2.
XX
XX 30-JUL-1998.
PD
XX
XX 22-JAN-1998; 98WO-US01216.
XX
XX 23-JAN-1997; 97US-0035917.
XX
XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX Kinzler KW, Velculescu VE, Vogelstein B;
PI
XX WPI: 1998-427943/36.
DR
XX
XX Yeast transcriptome - useful for modulating eukaryotic cell, for
PT screening antifungal agents, and for identifying genes in cell cycle
PT progression
XX
XX Claim 1; Page 25; 4App; English.
PS
XX
XX Yeast transcriptome is encoded by a DNA molecule comprising a yeast
CC gene involved in cell cycle progression selected from the group of
CC nonannotated ORF (NORF) genes. SAGE (serial analysis gene expression)
CC tags for highly expressed genes and NORF genes are given in AAV50051 to
CC AAV50345. The present invention describes: (1) a method of using yeast
CC genes to modulate the cell cycle which comprises administering to a cell
CC an isolated DNA molecule comprising a yeast gene which is involved in
CC cell cycle progression selected from differentially expressed genes
CC (SAGE tags given in AAV50051 to AAV50345); (2) a method for screening
CC candidate antifungal drugs which comprises contacting a test substance
CC with a yeast cell and monitoring expression of a yeast gene which is
CC involved in cell cycle progression; (3) a method of identifying human
CC genes which are involved in cell cycle progression which comprises
CC hybridizing a probe comprising at least 10 contiguous nucleotides of a
CC yeast gene which is differentially expressed between at least 2 phases
CC selected from the log phase, the S phase and the G2/M phase; and (4) a
CC probe for ascertaining the phase in the cell cycle, where the tags
CC comprises at least 14 contiguous nucleotides of a NORF gene (SAGE tags
CC given in AAV50051 to AAV50345), or as an array of probes on a solid
XX support.
XX
XX Sequence 10 BP; 5 A; 1 C; 1 G; 3 T; 0 other:
SO
XX
XX Query Match 100.0%; Score 6; DB 19; Length 10;
XX Best Local Similarity 83.3%; Pred. No. 1.6e+05;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
XX
XX 1 ATGACU 6
XX |||||:
XX 9 ATGACT 4
XX
XX
XX RESULT 10
XX AAX18647
XX AAX18647 standard; DNA; 10 BP.
XX
XX AAX18647;
XX
XX 06-MAY-1999 (first entry)
XX
XX p53 serial analysis of gene expression tag #50.
XX

```

```

KW p53; serial analysis of gene expression; SAGE tag; cancer; neoplastic;
KM rat embryo fibroblast; REF; tumour suppressor; cell cycle control;
XX tumorigenesis; diagnosis; ss.
OS Synthetic.
OS Rattus sp.
XX
XX W09901581-A1.
PN 14-JAN-1999.
XX
XX PD 02-JUL-1998; 98WO-US13903.
XX PE 02-JUL-1997; 97US-0051573.
XX PR 02-JUL-1997; 97US-0051573.
XX PA (GENZ ) GENZYME CORP.
XX PI Beaudry GA, Bertelsen AH, Galella EA, Madden SL,
XX WP1; 1999-106079/09.
XX DR
XX PT Diagnosis of cancer in potentially neoplastic samples - by comparing
PT the level of transcription between RNA transcripts in two tissue
PT samples, useful for providing an extensive profile of gene
PT expression in rat embryo fibroblast (REF) cells
XX
XX PS Example 2; Page 16; 32pp; English.
XX
XX A method has been developed for the diagnosis of cancer in potentially
CC neoplastic samples. The method comprises comparing the level of
CC transcription between RNA transcripts in two tissue samples (which are
CC of the same type), where the first sample is potentially neoplastic, and
CC the second sample is normal human tissue. The first sample is
CC categorized as neoplastic if its level of transcription is lower than
CC that of the second sample. The transcript is selected from Alu, RAS,
CC U6 snRNA, 16S rRNA, EGR-1, ribosomal protein S27, ETS-1, 28S rRNA, CGR11,
CC and LMK-2, and it is identified by a tag selected from ribosomal
CC protein l13a, alpha-tubulin (T1) and (T2), thymosin beta-4, and gamma-
CC actin. The present sequence represents a serial analysis of gene
CC expression (SAGE) tag from the present invention. The use of SAGE tags
CC provides an extensive profile of gene expression in rat embryo
CC fibroblast (REF) cells containing the (non)-functional p53 tumour
CC suppression gene. The discovery of new SAGE tags, which are regulated
CC by p53, enables the diagnosis of genes that are related to cell cycle
CC control and tumorigenesis.
XX
XX SO Sequence 10 BP; 2 A; 2 C; 3 G; 3 T; 0 other;
XX
XX Query Match 100.0%; Score 6; DB 20; Length 10;
XX Best Local Similarity 83.3%; Pred. No. 1.6e+05;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGACU 6
XX |||||:
Db 1 ATGACT 6
XX
XX RESULT 11
XX AAX18610/c
ID AAX18610 standard; DNA; 10 BP.
XX AC
XX AA18610;
XX DT 06-MAY-1999 (first entry)
XX
XX p53 serial analysis of gene expression tag #27.
DE
XX p53; serial analysis of gene expression; SAGE tag; cancer; neoplastic;
KM rat embryo fibroblast; REF; tumour suppressor; cell cycle control;
XX tumorigenesis; diagnosis; ss.
XX
XX Synthetic.
XX Rattus sp.

```

```

XX  W09901581-A1.
PN  14-JAN-1999.
XX  02-JUL-1998; 98WO-US13903.
XX  02-JUL-1997; 97US-0051573.
PR  (GENZ ) GENZYME CORP.
XX  Beauty GA, Bertelsen AH, Galella EA, Madden SR.
PI  WPI: 1999-106079/09.
XX  A method has been developed for the diagnosis of cancer in potentially
CC  neoplastic samples. The method comprises comparing the level of
CC  transcription between RNA transcripts in two tissue samples (which are
CC  of the same type), where the first sample is potentially neoplastic, and
CC  the second sample is normal human tissue. The first sample is
CC  categorized as neoplastic if its level of transcription is lower than
CC  that of the second sample. The transcript is selected from Alu, RAS,
CC  U6 snRNA, 16S RNA, EGR-1, ribosomal protein S27, ETS-1, 28S RNA, CGR11,
CC  and LIMK-2, and it is identified by a tag selected from ribosomal
CC  protein L13a, alpha-tubulin (T1) and (T2), thymosin beta-4, and gamma-
CC  actin. The present sequence represents a serial analysis of gene
CC  expression (SAGE) tag from the present invention. The use of SAGE tags
CC  provides an extensive profile of gene expression in rat embryo
CC  fibroblast (REF) cells containing the (non)-functional p53 tumour
CC  suppression gene. The discovery of new SAGE tags, which are regulated
CC  by p53, enables the diagnosis of genes that are related to cell cycle
CC  control and tumorigenesis.
XX  Sequence 10 BP; 2 A; 3 C; 2 G; 3 T; 0 other;
SQ
OY  1 ATGACU 6
   11111:
DB  8 ATGACT 3
   100.0%; Score 6; DB 20; Length 10;
   Best Local Similarity 83.3%; Pred. No. 1.6e+05;
   Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
AAAT6009
ID  AAAT6009 standard; DNA; 10 BP.
XX  AAAT6009:
XX  25-JAN-2001 (first entry)
XX  DNA sequence analysis oligonucleotide #6.
XX  DNA sequence analysis; stacking hybridisation; disease-associated allele;
XX  primer; ss.
XX  Synthetic.
XX  US6090549-A.
XX  18-JUL-2000.
XX  13-MAY-1997; 97US-0855372.
XX  16-JAN-1996; 96US-0587332.

```

```

XX  (UYCH-) UNIV CHICAGO.
PA  Parinov SV, Barski VE, Dubiley SA, Mirzabekov AD, Kirillov EV;
XX  WPI: 2000-523756/47.
XX  Determining number of repeat base sequences in a target oligonucleotide
XX  for diagnosing a disease, by detecting multiple mutation, utilizing
XX  continuous or contiguous stacking hybridization
XX  Example 3; Fig 6A; 35pp; English.
XX  The present sequence is a synthetic oligonucleotide which was used to
XX  demonstrate the methods of the invention. These involve analysing DNA
XX  sequences by hybridisation with oligonucleotides associated with
XX  polyacrylamide matrices, including continuous and/or contiguous stacking
XX  hybridisation. This can be used in sequencing, in diagnostic methods
XX  where different alleles are associated with a disease, and in sequencing
XX  long DNA fragments containing internal repeats, which enables the
XX  identification of unique sequences which may flank such repeats.
XX  Sequence 10 BP; 2 A; 1 C; 4 G; 3 T; 0 other;
SQ
OY  1 ATGACU 6
   11111:
DB  3 ATGACT 8
   100.0%; Score 6; DB 21; Length 10;
   Best Local Similarity 83.3%; Pred. No. 1.6e+05;
   Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
AAAT6025
ID  AAAT6025 standard; DNA; 10 BP.
XX  AAAT6025:
XX  25-JAN-2001 (first entry)
XX  DNA sequence analysis oligonucleotide #16.
XX  DNA sequence analysis; stacking hybridisation; disease-associated allele;
XX  primer; ss.
XX  Synthetic.
XX  US6090549-A.
XX  18-JUL-2000.
XX  13-MAY-1997; 97US-0855372.
XX  16-JAN-1996; 96US-0587332.
XX  (UYCH-) UNIV CHICAGO.
XX  Parinov SV, Barski VE, Dubiley SA, Mirzabekov AD, Kirillov EV;
XX  WPI: 2000-523756/47.
XX  Determining number of repeat base sequences in a target oligonucleotide
XX  for diagnosing a disease, by detecting multiple mutation, utilizing
XX  continuous or contiguous stacking hybridization
XX  Example 3; Column 12; 35pp; English.
XX  The present sequence is a synthetic oligonucleotide which was used to
XX  demonstrate the methods of the invention. These involve analysing DNA
XX  sequences by hybridisation with oligonucleotides associated with
XX  polyacrylamide matrices, including continuous and/or contiguous stacking
XX  hybridisation. This can be used in sequencing, in diagnostic methods
XX  where different alleles are associated with a disease, and in sequencing
XX  long DNA fragments containing internal repeats, which enables the
XX  identification of unique sequences which may flank such repeats.

```


CC where different alleles are associated with a disease, and in sequencing
 CC long DNA fragments containing internal repeats, which enables the
 CC identification of unique sequences which may flank such repeats.
 XX
 SQ Sequence 10 BP; 2 A; 1 C; 4 G; 3 T; 0 other;

Query Match 100.0%; Score 6; DB 21; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.6e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 DB 3 ATGACT 8

RESULT 14

AA277646
 ID AA277646 standard; DNA; 10 BP.

AC AA277646;

DT 10-APR-2000 (first entry)

DE Human dendritic cell SAGE tag, SEQ ID NO:74.

KW SAGE tag; serial analysis of gene expression; antigen-presenting cell;

KW APC; monocyte-derived dendritic cell; differential gene expression;

KW immunostimulatory cofactor; costimulatory factor; CTL;

KW cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.

OS Homo sapiens.

PN WO965924-A2.

PD 23-DEC-1999.

PF 18-JUN-1999;

PR 19-JUN-1998; 98US-0089833.

PR 19-JUN-1998; 98US-0089844.

PR 19-JUN-1998; 98US-0089853.

PR 19-JUN-1998; 98US-0089878.

PR 19-JUN-1998; 98US-0089911.

PR 19-JUN-1998; 98US-0089922.

PR 19-JUN-1998; 98US-0089933.

PR 19-JUN-1998; 98US-0089934.

PR 19-JUN-1998; 98US-0089937.

PR 19-JUN-1998; 98US-0089939.

PR 19-JUN-1998; 98US-0090000.

PR 19-JUN-1998; 98US-0090003.

PR 19-JUN-1998; 98US-0090036.

PR 19-JUN-1998; 98US-0090039.

PR 19-JUN-1998; 98US-0090040.

PR 19-JUN-1998; 98US-0090041.

PR 19-JUN-1998; 98US-0090042.

PR 19-JUN-1998; 98US-0090043.

PR 19-JUN-1998; 98US-0090044.

PR 19-JUN-1998; 98US-0090045.

PR 19-JUN-1998; 98US-0090047.

PR 19-JUN-1998; 98US-0090048.

PR 19-JUN-1998; 98US-0090072.

PR 19-JUN-1998; 98US-0090076.

PR 19-JUN-1998; 98US-0090077.

PR 19-JUN-1998; 98US-0090078.

PR 19-JUN-1998; 98US-0090079.

PR 19-JUN-1998; 98US-0090080.

PR 08-DEC-1998; 98US-0111715.

XX

XX

PA (GENZ) GENZYME CORP.

PA (ROBE/) ROBERTS B L.

PA (SHAN/) SHANKARA S.

XX

XX

PI Roberts BL, Shankara S;

XX WPI; 2000-106077/09.

PT Isolated polynucleotides differentially expressed in antigen-presenting
 PT cells, useful in gene vaccines against cancer

PS Claim 1; Page 65; 130pp; English.

CC Sequences AA277573-279709 represent SAGE (serial analysis of gene
 CC expression) tags used to identify mRNA transcripts encoding
 CC immunostimulatory cofactor proteins which are preferentially or
 CC differentially expressed in monocyte-derived dendritic cells compared
 CC with monocytes. Some of the transcripts correspond to known genes or
 CC ESTs (expressed sequence tags) which were previously unknown to be
 CC preferentially or differentially expressed in dendritic cells, while
 CC other transcripts correspond to novel genes. Antigen-presenting cell
 CC (APC)-associated costimulatory factors play an important role in the
 CC activation of the cytotoxic immune response, particularly against tumour
 CC cells. Tumour antigen presentation via the MHC (major histocompatibility
 CC complex) and subsequent recognition by T-cell receptors is alone
 CC insufficient to activate a robust cytotoxic immune response that can
 CC lyse the tumour cells, immunostimulatory cofactors also being required
 CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
 CC sequences identified using the SAGE tags have several potential uses.
 CC They may be used in vaccines to induce an immune response, particularly
 CC against a tumour antigen, to modulate the genotype of an APC, to screen
 CC for agents that modulate expression of differentially expressed genes in
 CC an APC; and as hybridisation probes/amplification primers for the
 CC diagnosis, prognosis and monitoring of diseases related to abnormal
 CC expression of these genes. Detection of the dendritic cell
 CC differentially expressed genes, or of their encoded proteins, can be used
 CC to identify cells as belonging to the monocyte lineage. Cells containing
 CC these genes can be used in active immunotherapy (or to stimulate
 CC production of a population of antigen-specific effector cells) and
 CC vectors containing them are used in gene therapy. Co-administration of
 CC tumour antigens and APC-associated costimulatory factors ensures adequate
 CC antigen presentation to endogenous APCs and upregulates the APCs for the
 CC presentation of co-stimulatory signals, migration to T cell-rich sites,
 CC secretion of T cell growth factors and secretion of chemokines for
 CC recruitment of immune effector cells.

SQ Sequence 10 BP; 4 A; 2 C; 2 G; 2 T; 0 other;

Query Match 100.0%; Score 6; DB 21; Length 10;

Best Local Similarity 83.3%; Pred. No. 1.6e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6

DB 5 ATGACT 10

RESULT 15

AA277660/c
 ID AA277660 standard; DNA; 10 BP.

AC AA277660;

DT 10-APR-2000 (first entry)

DE Human dendritic cell SAGE tag, SEQ ID NO:88.

KW SAGE tag; serial analysis of gene expression; antigen-presenting cell;

KW APC; monocyte-derived dendritic cell; differential gene expression;

KW immunostimulatory cofactor; costimulatory factor; CTL;

KW cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.

OS Homo sapiens.

PN WO965924-A2.

PD 23-DEC-1999.

XX

XX

XX

XX

XX

XX

XX

PF 18-JUN-1999; 99WO-US13800.
 XX 19-JUN-1998; 98US-0089833.
 PR 19-JUN-1998; 98US-0089844.
 PR 19-JUN-1998; 98US-0089853.
 PR 19-JUN-1998; 98US-0089878.
 PR 19-JUN-1998; 98US-0089911.
 PR 19-JUN-1998; 98US-0089921.
 PR 19-JUN-1998; 98US-0089923.
 PR 19-JUN-1998; 98US-0089924.
 PR 19-JUN-1998; 98US-0089927.
 PR 19-JUN-1998; 98US-0089929.
 PR 19-JUN-1998; 98US-0090000.
 PR 19-JUN-1998; 98US-0090035.
 PR 19-JUN-1998; 98US-0090036.
 PR 19-JUN-1998; 98US-0090039.
 PR 19-JUN-1998; 98US-0090040.
 PR 19-JUN-1998; 98US-0090041.
 PR 19-JUN-1998; 98US-0090042.
 PR 19-JUN-1998; 98US-0090043.
 PR 19-JUN-1998; 98US-0090044.
 PR 19-JUN-1998; 98US-0090045.
 PR 19-JUN-1998; 98US-0090047.
 PR 19-JUN-1998; 98US-0090048.
 PR 19-JUN-1998; 98US-0090072.
 PR 19-JUN-1998; 98US-0090076.
 PR 19-JUN-1998; 98US-0090077.
 PR 19-JUN-1998; 98US-0090078.
 PR 19-JUN-1998; 98US-0090079.
 PR 19-JUN-1998; 98US-0090080.
 PR 08-DEC-1998; 98US-0111715.

XX (GENZ) GENZYME CORP.
 PA (ROBE/) ROBERTS B L.
 PA (SHAN/) SHANKARA S.

PI Roberts BL, Shankara S;

DR WPI; 2000-106077/09.

XX Isolated polynucleotides differentially expressed in antigen-presenting
 PT cells, useful in gene vaccines against cancer -

PS Claim 1; Page 66; 130pp; English.

XX Sequences AA277573-279709 represent SAGE (serial analysis of gene
 CC expression) tags used to identify mRNA transcripts encoding
 CC immunostimulatory cofactor proteins which are preferentially or
 CC differentially expressed in monocyte-derived dendritic cells compared
 CC with monocytes. Some of the transcripts correspond to known genes or
 CC ESTs (expressed sequence tags) which were previously unknown to be
 CC preferentially or differentially expressed in dendritic cells, while
 CC other transcripts correspond to novel genes. Antigen-presenting cell
 CC (APC)-associated costimulatory factors play an important role in the
 CC activation of the cytotoxic immune response, particularly against tumour
 CC cells. Tumour antigen presentation via the MHC (major histocompatibility
 CC complex) and subsequent recognition by T-cell receptors is alone
 CC insufficient to activate a robust cytotoxic immune response that can
 CC lyse the tumour cells. Immunostimulatory cofactors also being required
 CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
 CC sequences identified using the SAGE tags have several potential uses.
 CC They may be used in vaccines to induce an immune response, particularly
 CC against a tumour antigen; to modulate the genotype of an APC; to screen
 CC for agents that modulate expression of differentially expressed genes in
 CC an APC; and as hybridisation probes/amplification primers for the
 CC diagnosis, prognosis and monitoring of diseases related to abnormal
 CC expression of these genes. Detection of the dendritic cell
 CC differentially expressed genes, or of their encoded proteins, can be used
 CC to identify cells as belonging to the monocyte lineage. Cells containing
 CC these genes can be used in active immunotherapy (or to stimulate
 CC production of a population of antigen-specific effector cells) and
 CC vectors containing them are used in gene therapy. Co-administration of
 CC tumour antigens and APC-associated costimulatory factors ensures adequate

CC antigen presentation to endogenous APCs and upregulates the APCs for the
 CC presentation of co-stimulatory signals, migration to T cell-rich sites,
 CC secretion of T cell growth factors and secretion of chemokines for
 CC recruitment of immune effector cells.

XX Sequence 10 BP; 2 A; 2 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 6; DB 21; Length 10;

Best Local Similarity 83.3%; Pred. No. 1.6e+05; Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX
 OY 1 ATGACU 6
 DB 7 ATGACT 2

Search completed: May 21, 2003, 04:50:10
 Job time : 159 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:37:42 ; Search time 35.25 Seconds
(Without alignments)
52,200 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6

Sequence: 1 atgacg 6

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cg2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cg2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cg2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cg2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cg2_6/ptodata/2/ina/PCUTS.COMB.seq:*
6: /cg2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|----|--------------------|
| 1 | 6 | 100.0 | 8 | US-08-859-954-66 |
| 2 | 6 | 100.0 | 8 | US-08-859-954-158 |
| 3 | 6 | 100.0 | 10 | US-08-388-353-232 |
| 4 | 6 | 100.0 | 10 | US-08-388-353-233 |
| 5 | 6 | 100.0 | 10 | US-08-388-353-234 |
| 6 | 6 | 100.0 | 10 | US-08-388-353-235 |
| 7 | 6 | 100.0 | 10 | US-08-388-353-236 |
| 8 | 6 | 100.0 | 10 | US-08-488-551B-232 |
| 9 | 6 | 100.0 | 10 | US-08-488-551B-233 |
| 10 | 6 | 100.0 | 10 | US-08-488-551B-234 |
| 11 | 6 | 100.0 | 10 | US-08-488-551B-235 |
| 12 | 6 | 100.0 | 10 | US-08-488-551B-236 |
| 13 | 6 | 100.0 | 10 | US-08-855-372B-19 |
| 14 | 6 | 100.0 | 10 | US-08-855-372B-87 |
| 15 | 6 | 100.0 | 10 | US-08-991-552B-87 |
| 16 | 6 | 100.0 | 10 | US-09-498-851-19 |
| 17 | 6 | 100.0 | 10 | US-09-498-851-87 |
| 18 | 6 | 100.0 | 11 | US-08-242-664-10 |
| 19 | 6 | 100.0 | 11 | US-08-467-219-3 |
| 20 | 6 | 100.0 | 11 | US-08-484-138-10 |
| 21 | 6 | 100.0 | 11 | US-08-659-924-3 |
| 22 | 6 | 100.0 | 11 | US-09-196-523-10 |
| 23 | 6 | 100.0 | 11 | PCT-US95-06379-10 |
| 24 | 6 | 100.0 | 12 | US-08-441-887A-154 |
| 25 | 6 | 100.0 | 12 | US-08-441-887A-155 |
| 26 | 6 | 100.0 | 12 | US-09-281-418-139 |
| 27 | 6 | 100.0 | 12 | US-08-532-657A-2 |

ALIGNMENTS

| | | | | | | |
|----|---|-------|----|---|-------------------|-------------------|
| 28 | 6 | 100.0 | 12 | 4 | US-09-631-349A-12 | Sequence 12, Appl |
| 29 | 6 | 100.0 | 13 | 2 | US-08-867-820A-49 | Sequence 49, Appl |
| 30 | 6 | 100.0 | 13 | 6 | US-08-797-812-16 | Sequence 16, Appl |
| 31 | 6 | 100.0 | 13 | 6 | 522537-13 | Patent No. 522537 |
| 32 | 6 | 100.0 | 14 | 1 | US-08-303-004-1 | Sequence 1, Appl |
| 33 | 6 | 100.0 | 14 | 1 | US-08-358-810A-3 | Sequence 3, Appl |
| 34 | 6 | 100.0 | 14 | 1 | US-08-484-712A-3 | Sequence 3, Appl |
| 35 | 6 | 100.0 | 14 | 1 | US-08-359-295C-16 | Sequence 16, Appl |
| 36 | 6 | 100.0 | 14 | 2 | US-08-485-105A-16 | Sequence 16, Appl |
| 37 | 6 | 100.0 | 14 | 2 | US-08-769-945C-2 | Sequence 2, Appl |
| 38 | 6 | 100.0 | 14 | 2 | US-08-769-945C-3 | Sequence 3, Appl |
| 39 | 6 | 100.0 | 14 | 2 | US-08-769-945C-4 | Sequence 4, Appl |
| 40 | 6 | 100.0 | 14 | 2 | US-08-769-945C-15 | Sequence 15, Appl |
| 41 | 6 | 100.0 | 14 | 2 | US-08-769-945C-17 | Sequence 17, Appl |
| 42 | 6 | 100.0 | 14 | 2 | US-08-769-945C-18 | Sequence 18, Appl |
| 43 | 6 | 100.0 | 14 | 2 | US-08-769-945C-19 | Sequence 19, Appl |
| 44 | 6 | 100.0 | 14 | 3 | US-09-183-650-16 | Sequence 16, Appl |
| 45 | 6 | 100.0 | 14 | 4 | US-09-081-646-7 | Sequence 7, Appl |

RESULT 1
US-08-859-954-66/c
Sequence 66, Application US/08859954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramon
TITLE OF INVENTION: Design and Optimized Primer Library for
TITLE OF INVENTION: Gene Sequencing and Method thereof
NUMBER OF SEQUENCES: 566
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,954
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ. ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligonucleotide"
HYPOTHETICAL: YES
ANTI-SENSE: YES
US-08-859-954-66
Query Match 100.0%, Score 6, DB 3, Length 8;

Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 8 ATGACT 3

RESULT 2

US-08-859-954-158
; Sequence 158, Application US/08859954
; Patent No. 6083695
; GENERAL INFORMATION:
; APPLICANT: Hardin, Susan H.
; APPLICANT: Homayouni, Ramin
; APPLICANT: Hardin, Paul E.
; TITLE OF INVENTION: Design and Optimized Primer Library for
; TITLE OF INVENTION: Gene Sequencing and Method Thereof
; NUMBER OF SEQUENCES: 566
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,954
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,782
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5900
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5325
; TELEFAX: 713/651-5246
; INFORMATION FOR SEQ ID NO: 158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "oligonucleotide"
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; US-08-859-954-158

Query Match 100.0%; Score 6; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 2 ATGACT 7

RESULT 3

US-08-388-353-232
; Sequence 232, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.

; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 232:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; US-08-388-353-232

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
Db 5 ATGACT 10

RESULT 4

US-08-388-353-233
; Sequence 233, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353

FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 233:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-233

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGACU 6
|||||
Db 4 ATGACT 9

RESULT 5
US-08-388-353-234
Sequence 234, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 234:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-234

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGACU 6
|||||
Db 3 ATGACT 8

RESULT 6
US-08-388-353-235
Sequence 235, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 235:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-235

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGACU 6
|||||
Db 2 ATGACT 7

RESULT 7
US-08-388-353-236
Sequence 236, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David

;; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
;; NUMBER OF SEQUENCES: 800
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Scully, Scott, Murphy & Presser
;; STREET: 400 Garden City Plaza
;; CITY: Garden City
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 11530
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/388,353
;; FILING DATE: 14-FEB-1995
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: DIGILLO, Frank S.
;; REGISTRATION NUMBER: 31,346
;; REFERENCE/DOCKET NUMBER: 9606
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 742-4343
;; TELEFAX: (516) 742-4366
;; TELEX: 230 901 SANS UR
;; INFORMATION FOR SEQ ID NO: 236:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-388-353-236

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
DB 1 ATGACT 6

RESULT 8

;; Sequence 232, Application US/08488551B
;; Patent No. 6015661
;; GENERAL INFORMATION:
;; APPLICANT: Nicholas J. Deacon
;; APPLICANT: Dale A. McPhee
;; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
;; NUMBER OF SEQUENCES: 841
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;; STREET: 400 GARDEN CITY PLAZA
;; CITY: GARDEN CITY
;; STATE: NEW YORK
;; COUNTRY: U.S.A.
;; ZIP: 11530-0299
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/488,551B
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PM3864 (AU)
;; FILING DATE: 14-FEB-1994
;; APPLICATION NUMBER: PM4002 (AU)

;; FILING DATE: 21-FEB-1994
;; APPLICATION NUMBER: PM0284 (AU)
;; FILING DATE: 23-DEC-1994
;; APPLICATION NUMBER: US 08/388,353
;; FILING DATE: 14-FEB-1995
;; APPLICATION NUMBER: PM3021/95
;; FILING DATE: 17-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FRANK S. DIGILLO
;; REFERENCE/DOCKET NUMBER: 96062
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 742-4343
;; TELEFAX: (516) 742-4366
;; INFORMATION FOR SEQ ID NO: 232:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-488-551B-232

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
DB 5 ATGACT 10

RESULT 9

;; Sequence 233, Application US/08488551B
;; Patent No. 6015661
;; GENERAL INFORMATION:
;; APPLICANT: Nicholas J. Deacon
;; APPLICANT: Dale A. McPhee
;; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
;; NUMBER OF SEQUENCES: 841
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;; STREET: 400 GARDEN CITY PLAZA
;; CITY: GARDEN CITY
;; STATE: NEW YORK
;; COUNTRY: U.S.A.
;; ZIP: 11530-0299
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/488,551B
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PM3864 (AU)
;; FILING DATE: 14-FEB-1994
;; APPLICATION NUMBER: PM4002 (AU)
;; FILING DATE: 21-FEB-1994
;; APPLICATION NUMBER: PM0284 (AU)
;; FILING DATE: 23-DEC-1994
;; APPLICATION NUMBER: US 08/388,353
;; FILING DATE: 14-FEB-1995
;; APPLICATION NUMBER: PM3021/95
;; FILING DATE: 17-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FRANK S. DIGILLO
;; REFERENCE/DOCKET NUMBER: 96062
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 742-4343
;; TELEFAX: (516) 742-4366

INFORMATION FOR SEQ ID NO: 233:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-233

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 4 ATGACT 9

RESULT 10
US-08-488-551B-234
Sequence 234, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 9606Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 234:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-234

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 3 ATGACT 8

RESULT 11
US-08-488-551B-235
Sequence 235, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 9606Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 235:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-235

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 2 ATGACT 7

RESULT 12
US-08-488-551B-236
Sequence 236, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper

TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-236

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
Db 1 ATGACT 6

RESULT 13
US-08-855-372B-19
Sequence 19, Application US/08855372B
Patent No. 6090549
GENERAL INFORMATION:
APPLICANT: Mirzabekov, Andrei D
APPLICANT: Parinov, Sergei V
APPLICANT: Barsky, Victor E
APPLICANT: Kirillov, Eugene V
APPLICANT: Dubilley, Svetlana A
TITLE OF INVENTION: Use of Continuous/Contiguous Stacking Hybridization as a Diag
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHERSKOV & FLAYNIK
STREET: 20 N. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch, 1.4 MB storage

COMPUTER: PC
OPERATING SYSTEM: Microsoft Windows 98
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,372B
FILING DATE: 13-MAY-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/587,332
FILING DATE: 16-JAN-96
ATTORNEY/AGENT INFORMATION:
NAME: Cherskov, Michael J.
REGISTRATION NUMBER: 33,664
REFERENCE/DOCKET NUMBER: ANL-IN-95-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 621-1330
TELEFAX: (312) 621-0088
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 bases
TYPE: nucleic acid
STRANDEDNESS: No. 6090549 Applicable
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: yes
US-08-855-372B-19

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
Db 3 ATGACT 8

RESULT 14
US-08-855-372B-87
Sequence 87, Application US/08855372B
Patent No. 6090549
GENERAL INFORMATION:
APPLICANT: Mirzabekov, Andrei D
APPLICANT: Parinov, Sergei V
APPLICANT: Barsky, Victor E
APPLICANT: Kirillov, Eugene V
APPLICANT: Dubilley, Svetlana A
TITLE OF INVENTION: Use of Continuous/Contiguous Stacking Hybridization as a D1
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHERSKOV & FLAYNIK
STREET: 20 N. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch, 1.4 MB storage
COMPUTER: PC
OPERATING SYSTEM: Microsoft Windows 98
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,372B
FILING DATE: 13-MAY-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/587,332
FILING DATE: 16-JAN-96
ATTORNEY/AGENT INFORMATION:
NAME: Cherskov, Michael J.
REGISTRATION NUMBER: 33,664
REFERENCE/DOCKET NUMBER: ANL-IN-95-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 621-1330
TELEFAX: (312) 621-0088
INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:
 LENGTH: 10 bases
 TYPE: nucleic acid
 STRANDEDNESS: No. 6090549 Applicable
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 HYPOTHETICAL: yes
 US-08-855-372B-87

Query Match 100.0%; Score 6; DB 3; Length 10;
 Best Local Similarity 83.3%; Pred. No. 2e+04;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 3 ATGACT 8

RESULT 15
 US-08-991-525B-87
 ; Sequence 87, Application US/08991525B
 ; Patent No. 6093811
 ; GENERAL INFORMATION:
 ; APPLICANT: Bennett and Mirabelli
 ; TITLE OF INVENTION: Oligonucleotide Modulation
 ; TITLE OF INVENTION: of Cell Adhesion
 ; NUMBER OF SEQUENCES: 87
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Offices of Jane Massey Licata
 ; STREET: 66 East Main Street
 ; CITY: Marlton
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 08053
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WORDPERFECT 6.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/991,525B
 ; FILING DATE: December 16, 1997
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 440,740
 ; FILING DATE: May 12, 1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 063,167
 ; FILING DATE: May 17, 1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 969,151
 ; FILING DATE: February 10, 1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 007,997
 ; FILING DATE: January 21, 1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 939,855
 ; FILING DATE: September 2, 1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 567,286
 ; FILING DATE: August 14, 1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jane Massey Licata
 ; REGISTRATION NUMBER: 32,257
 ; REFERENCE/DOCKET NUMBER: ISPH-0247
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (856) 810-1515
 ; TELEFAX: (856) 810-1454
 ; INFORMATION FOR SEQ ID NO: 87:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 10
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single

TOPOLOGY: linear
 ; ANTI-SENSE: yes
 US-08-991-525B-87

Query Match 100.0%; Score 6; DB 3; Length 10;
 Best Local Similarity 83.3%; Pred. No. 2e+04;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
 |||||
 Db 5 ATGACT 10

Search completed: May 21, 2003, 06:28:34
 Job time : 38.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:39:53 ; Search time 238 Seconds
(without alignments)
33.289 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6
Sequence: 1 atgacu 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTC_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1 | 6 | 100.0 | 9 | US-09-728-574-14 | Sequence 14, Appl |
| 2 | 6 | 100.0 | 10 | US-10-010-802-278 | Sequence 278, Appl |
| 3 | 6 | 100.0 | 10 | US-09-488-220-18 | Sequence 18, Appl |
| 4 | 6 | 100.0 | 10 | US-10-033-145-74 | Sequence 74, Appl |
| 5 | 6 | 100.0 | 10 | US-10-033-145-88 | Sequence 88, Appl |
| 6 | 6 | 100.0 | 10 | US-10-033-145-98 | Sequence 98, Appl |
| 7 | 6 | 100.0 | 10 | US-10-033-145-795 | Sequence 795, Appl |
| 8 | 6 | 100.0 | 10 | US-10-033-145-1256 | Sequence 1256, Ap |
| 9 | 6 | 100.0 | 10 | US-10-033-145-2043 | Sequence 2043, Ap |
| 10 | 6 | 100.0 | 11 | US-09-249-135-16 | Sequence 16, Appl |
| 11 | 6 | 100.0 | 13 | US-09-818-875-4349 | Sequence 4349, Ap |
| 12 | 6 | 100.0 | 14 | US-08-591-486B-13 | Sequence 13, Appl |
| 13 | 6 | 100.0 | 14 | US-08-591-486B-48 | Sequence 48, Appl |
| 14 | 6 | 100.0 | 14 | US-09-998-027-41 | Sequence 41, Appl |
| 15 | 6 | 100.0 | 15 | US-09-880-887-10 | Sequence 10, Appl |
| 16 | 6 | 100.0 | 15 | US-09-867-915-10 | Sequence 10, Appl |
| 17 | 6 | 100.0 | 15 | US-10-082-476-2 | Sequence 2, Appl |
| 18 | 6 | 100.0 | 15 | US-10-056-414-118 | Sequence 118, Appl |
| 19 | 6 | 100.0 | 15 | US-10-043-875-181 | Sequence 181, Appl |

| | | | | | | |
|------|---|-------|----|----|---------------------|--------------------|
| c 20 | 6 | 100.0 | 15 | 9 | US-09-848-754A-9112 | Sequence 9112, Ap |
| c 21 | 6 | 100.0 | 15 | 9 | US-10-010-802-50 | Sequence 50, Appl |
| c 22 | 6 | 100.0 | 15 | 9 | US-10-010-802-51 | Sequence 51, Appl |
| c 23 | 6 | 100.0 | 15 | 9 | US-10-010-802-172 | Sequence 172, Appl |
| c 24 | 6 | 100.0 | 15 | 9 | US-10-010-802-174 | Sequence 174, Appl |
| c 25 | 6 | 100.0 | 15 | 9 | US-10-010-802-176 | Sequence 176, Appl |
| c 26 | 6 | 100.0 | 15 | 9 | US-10-010-802-178 | Sequence 178, Appl |
| c 27 | 6 | 100.0 | 15 | 9 | US-09-439-429-19 | Sequence 19, Appl |
| c 28 | 6 | 100.0 | 15 | 9 | US-10-287-919-2017 | Sequence 2017, Ap |
| c 29 | 6 | 100.0 | 15 | 10 | US-09-504-231A-481 | Sequence 481, App |
| c 30 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1100 | Sequence 1100, Ap |
| c 31 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1220 | Sequence 1220, Ap |
| c 32 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1221 | Sequence 1221, Ap |
| c 33 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1242 | Sequence 1242, Ap |
| c 34 | 6 | 100.0 | 15 | 10 | US-09-179-536B-295 | Sequence 295, Appl |
| c 35 | 6 | 100.0 | 15 | 10 | US-09-780-954A-9 | Sequence 9, Appl |
| c 36 | 6 | 100.0 | 15 | 10 | US-09-780-954A-10 | Sequence 10, Appl |
| c 37 | 6 | 100.0 | 15 | 10 | US-09-274-553D-481 | Sequence 481, App |
| c 38 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1100 | Sequence 1100, Ap |
| c 39 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1220 | Sequence 1220, Ap |
| c 40 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1221 | Sequence 1221, Ap |
| c 41 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1242 | Sequence 1242, Ap |
| c 42 | 6 | 100.0 | 15 | 10 | US-09-953-242-4 | Sequence 4, Appl |
| c 43 | 6 | 100.0 | 15 | 10 | US-09-953-242-5 | Sequence 5, Appl |
| c 44 | 6 | 100.0 | 15 | 10 | US-09-953-242-13 | Sequence 13, Appl |
| c 45 | 6 | 100.0 | 16 | 8 | US-08-591-486B-11 | Sequence 11, Appl |

ALIGNMENTS

RESULT 1
US-09-728-574-14
; Sequence 14, Application US/09728574
; Patent No. US20020137036A1
; GENERAL INFORMATION:
; APPLICANT: Strategene
; TITLE OF INVENTION: Methods for Detection of a Target Nucleic Acid By Capture
; FILE REFERENCE: 25436/1660
; CURRENT APPLICATION NUMBER: US/09/728,574
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 09/728574
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Saccharomyces sp.
; FEATURE:
; NAME/KEY: GCN4 DNA binding site
; LOCATOR: (1)..(9)
US-09-728-574-14

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 1 ATGACT 6

RESULT 2
US-10-010-802-278/c
; Sequence 278, Application US/10010802
; Publication No. US20030078220A1
; GENERAL INFORMATION:
; APPLICANT: Genalsance Pharmaceuticals
; APPLICANT: Chew, Anne
; APPLICANT: Denton, R. Rex
; APPLICANT: Duda, Amy
; APPLICANT: Nandabalan, Krishnan

APPLICANT: Stephens, J. Claiborne
APPLICANT: Windemuth, Andreas
TITLE OF INVENTION: Drug Target Isogenes: Polymorphisms in the Interleukin
TITLE OF INVENTION: 4 Receptor Alpha Gene
FILE REFERENCE: MMH-0002US2 ILAR alpha
CURRENT APPLICATION NUMBER: US/10/010,802
CURRENT FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: PCT/US00/19094
PRIOR FILING DATE: 2000-07-13
NUMBER OF SEQ ID NOS: 413
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 278
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-010-802-278

Query Match 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 9 ATGACT 4

RESULT 3

US-09-489-220-18
Sequence 18, Application US/09489220
Patent No. US20020110808A1
GENERAL INFORMATION:
APPLICANT: Reidhaer-Olson, John F.
TITLE OF INVENTION: Toxicant-Induced Differential Gene Expression
FILE REFERENCE: 16528A-038900US
CURRENT APPLICATION NUMBER: US/09/489,220
CURRENT FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ten nucleotides
OTHER INFORMATION: following M13R(-48) primer sequence in 5'
OTHER INFORMATION: arbitrary primer (Arp) 14
US-09-489-220-18

Query Match 100.0%; Score 6; DB 10; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 4 ATGACT 9

RESULT 4

US-10-033-145-74
Sequence 74, Application US/10033145
Patent No. US2002015151A1
GENERAL INFORMATION:
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GA0201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137

SOFTWARE: PatentIn version 3.0
SEQ ID NO 74
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-74

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 5 ATGACT 10

RESULT 5

US-10-033-145-88/C
Sequence 88, Application US/10033145
Patent No. US2002015151A1
GENERAL INFORMATION:
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GA0201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137
SOFTWARE: PatentIn version 3.0
SEQ ID NO 88
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-88

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 7 ATGACT 2

RESULT 6

US-10-033-145-98
Sequence 98, Application US/10033145
Patent No. US2002015151A1
GENERAL INFORMATION:
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GA0201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137
SOFTWARE: PatentIn version 3.0
SEQ ID NO 98
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-98

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
DB 2 ATGACT 7

RESULT 7
US-10-033-145-795/c
; Sequence 795, Application US/10033145
; Patent No. US2002015151A1

GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION

APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKAR, SRINIVAS

TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES

FILE REFERENCE: GAO201C

CURRENT APPLICATION NUMBER: US/10/033,145

CURRENT FILING DATE: 2001-11-05

PRIOR APPLICATION NUMBER: PCT/US99/13800

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 2137

SOFTWARE: PatentIn version 3.0

SEQ ID NO 795

LENGTH: 10

TYPE: DNA

ORGANISM: Homo sapiens

US-10-033-145-795

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
DB 8 ATGACT 3

RESULT 8
US-10-033-145-1256
; Sequence 1256, Application US/10033145
; Patent No. US2002015151A1

GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION

APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKAR, SRINIVAS

TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES

FILE REFERENCE: GAO201C

CURRENT APPLICATION NUMBER: US/10/033,145

CURRENT FILING DATE: 2001-11-05

PRIOR APPLICATION NUMBER: PCT/US99/13800

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 2137

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1256

LENGTH: 10

TYPE: DNA

ORGANISM: Homo sapiens

US-10-033-145-1256

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
DB 2 ATGACT 7

RESULT 9
US-10-033-145-2043
; Sequence 2043, Application US/10033145
; Patent No. US2002015151A1

GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION

APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKAR, SRINIVAS

APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKAR, SRINIVAS

TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES

FILE REFERENCE: GAO201C

CURRENT APPLICATION NUMBER: US/10/033,145

CURRENT FILING DATE: 2001-11-05

PRIOR APPLICATION NUMBER: PCT/US99/13800

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 2137

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2043

LENGTH: 10

TYPE: DNA

ORGANISM: Homo sapiens

US-10-033-145-2043

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
DB 1 ATGACT 6

RESULT 10
US-09-249-155-16

; Sequence 16, Application US/09249155
; Publication No. US20030037345A1

GENERAL INFORMATION:
; APPLICANT: Heber-Katz, Ellen

TITLE OF INVENTION: Compositions and Methods for Wound

FILE REFERENCE: 00486.78503

CURRENT APPLICATION NUMBER: US/09/249,155

CURRENT FILING DATE: 1999-02-12

EARLIER APPLICATION NUMBER: 60/074,737

EARLIER FILING DATE: 1998-02-13

EARLIER APPLICATION NUMBER: 60/097,937

EARLIER FILING DATE: 1998-08-26

EARLIER APPLICATION NUMBER: 60/102,051

EARLIER FILING DATE: 1998-09-28

NUMBER OF SEQ ID NOS: 254

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 16

LENGTH: 11

TYPE: DNA

ORGANISM: Mus musculus

US-09-249-155-16

Query Match 100.0%; Score 6; DB 9; Length 11;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
DB 1 ATGACT 6

RESULT 11
US-09-818-875-4349

; Sequence 4349, Application US/09818875
; Publication No. US20030051270A1

GENERAL INFORMATION:
; APPLICANT: Knieg, Eric B.

APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.

TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single

FILE REFERENCE: Napro-4

CURRENT APPLICATION NUMBER: US/09/818,875

CURRENT FILING DATE: 2001-03-27

PRIOR APPLICATION NUMBER: US 60/192,176

PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/192,179
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/208,538
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/244,989
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 4385
SOFTWARE: Friedmann macro Naprod
SEQ ID NO 4349
LENGTH: 13
TYPE: DNA
ORGANISM: Escherichia coli
US-09-818-875-4349

Query Match
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
DB 8 ATGACT 13

RESULT 12
US-08-591-486B-13/c
Sequence 13, Application US/08591486B
Patent No. US20020037866A1
GENERAL INFORMATION:
APPLICANT: Schlingensiepen, Georg F
APPLICANT: Schlingensiepen, Reimar
APPLICANT: Gottingen, Wolfgang Brysch
TITLE OF INVENTION: A Pharmaceutical Composition
TITLE OF INVENTION: of Neuronal Injury, Degeneration and Cell Death and for the
TITLE OF INVENTION: Treatment of Neoplasms
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern
STREET: 400 Seventh Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,486B
FILING DATE: 11-JAN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93111059.7
FILING DATE: 10-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02218
FILING DATE: 6-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: 10496/P60122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-9350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: YES
US-08-591-486B-13

Query Match
Best Local Similarity 100.0%; Score 6; DB 8; Length 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
DB 9 ATGACT 4

RESULT 13
US-08-591-486B-48
Sequence 48, Application US/08591486B
Patent No. US20020037866A1
GENERAL INFORMATION:
APPLICANT: Schlingensiepen, Georg F
APPLICANT: Schlingensiepen, Reimar
APPLICANT: Gottingen, Wolfgang Brysch
TITLE OF INVENTION: A Pharmaceutical Composition
TITLE OF INVENTION: of Neuronal Injury, Degeneration and Cell Death and for the
TITLE OF INVENTION: Treatment of Neoplasms
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern
STREET: 400 Seventh Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,486B
FILING DATE: 11-JAN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93111059.7
FILING DATE: 10-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02218
FILING DATE: 6-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: 10496/P60122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-9350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: YES
US-08-591-486B-48

Query Match
Best Local Similarity 100.0%; Score 6; DB 8; Length 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
DB 9 ATGACT 14

RESULT 14

US-09-998-027-41/c
; Sequence 41, Application US/09998027
; Publication No. US20030093819A1
; GENERAL INFORMATION:
; APPLICANT: D/Andrea et al.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Diagnosis and Treatment of Cancers Associated with Defective
; TITLE OF INVENTION: DNA Repair Mechanisms
; FILE REFERENCE: 2486/101
; CURRENT APPLICATION NUMBER: US/09/998,027
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; OTHER INFORMATION: Intron/Exon Junctions of FANCD
US-09-998-027-41

Query Match

100.0%; Score 6; DB 9; Length 14;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||

Db 12 ATGACT 7

RESULT 15

US-09-880-887-10/c
; Sequence 10, Application US/09880887
; Patent No. US20020165177A1
; GENERAL INFORMATION:
; APPLICANT: NEGRIER, CLAUDE
; APPLICANT: PLANTIER, JEAN LUC
; TITLE OF INVENTION: MODIFIED FACTOR VIII CDNA
; FILE REFERENCE: 06478.1441
; CURRENT APPLICATION NUMBER: US/09/880,887
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 09/526,935
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: EP 99104050.2
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
US-09-880-887-10

Query Match

100.0%; Score 6; DB 9; Length 15;
Best Local Similarity 83.3%; Pred. No. 1.3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||

Db 8 ATGACT 3

Search completed: May 21, 2003, 06:44:43
Job time : 241 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:13:38 ; Search time 1094 seconds
(without alignments)
88.824 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6
Sequence: 1 atgacg 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 6 | 100.0 | 19 | 9 | AA885444 aml14h07.s |
| 2 | 6 | 100.0 | 19 | 13 | BM395830 5009-0-12 |
| 3 | 6 | 100.0 | 19 | 13 | BM396331 5009-0-2- |
| 4 | 6 | 100.0 | 19 | 13 | BM399859 5009-0-62 |
| 5 | 6 | 100.0 | 19 | 14 | D44776 HUMSUPY214 |
| 6 | 6 | 100.0 | 19 | 17 | AZ317020 1M0035P18 |

| | | | | | | |
|----|---|-------|----|----|-----------|---------------------|
| 7 | 6 | 100.0 | 19 | 17 | AZ457710 | AZ457710 1M0261K07 |
| 8 | 6 | 100.0 | 19 | 17 | AZ658087 | AZ658087 1M0534J09 |
| 9 | 6 | 100.0 | 19 | 17 | AZ653240 | AZ653240 1M0542H18 |
| 10 | 6 | 100.0 | 19 | 17 | AZ763729 | AZ763729 1M0559N16 |
| 11 | 6 | 100.0 | 19 | 17 | AZ861634 | AZ861634 2M0168B18 |
| 12 | 6 | 100.0 | 19 | 17 | AZ990851 | AZ990851 2M0274E15 |
| 13 | 6 | 100.0 | 20 | 17 | AZ500712 | AZ500712 1M0339B13 |
| 14 | 6 | 100.0 | 20 | 17 | AZ830285 | AZ830285 2M0109N22 |
| 15 | 6 | 100.0 | 21 | 17 | AZ455887 | AZ455887 1M0258019 |
| 16 | 6 | 100.0 | 21 | 17 | AZ662959 | AZ662959 1M0542P18 |
| 17 | 6 | 100.0 | 21 | 17 | AZ806895 | AZ806895 2M0069E13 |
| 18 | 6 | 100.0 | 21 | 17 | AZ949072 | AZ949072 2M0212D19 |
| 19 | 6 | 100.0 | 22 | 9 | A1001081 | A1001081 0566409.s |
| 20 | 6 | 100.0 | 22 | 17 | AZ303795 | AZ303795 1M00033107 |
| 21 | 6 | 100.0 | 22 | 17 | AZ314083 | AZ314083 1M0030P11 |
| 22 | 6 | 100.0 | 22 | 17 | AZ394940 | AZ394940 1M0158P11 |
| 23 | 6 | 100.0 | 22 | 17 | AZ651343 | AZ651343 1M0522D06 |
| 24 | 6 | 100.0 | 22 | 17 | AZ829109 | AZ829109 2M0106L03 |
| 25 | 6 | 100.0 | 22 | 17 | BH863544 | BH863544 SALK 0940 |
| 26 | 6 | 100.0 | 22 | 17 | TA285H08Q | AL485478 T. brucei |
| 27 | 6 | 100.0 | 23 | 13 | BM395136 | BM395136 50072-2-7 |
| 28 | 6 | 100.0 | 23 | 13 | BM396053 | BM396053 5009-0-16 |
| 29 | 6 | 100.0 | 23 | 17 | AZ307676 | AZ307676 1M0009N24 |
| 30 | 6 | 100.0 | 23 | 17 | AZ647964 | AZ647964 1M0514F13 |
| 31 | 6 | 100.0 | 23 | 17 | AZ808094 | AZ808094 2M0071K24 |
| 32 | 6 | 100.0 | 23 | 17 | AZ824304 | AZ824304 2M0098N17 |
| 33 | 6 | 100.0 | 23 | 17 | AZ838947 | AZ838947 2M0134B21 |
| 34 | 6 | 100.0 | 23 | 17 | BH790464 | BH790464 SALK 0571 |
| 35 | 6 | 100.0 | 23 | 17 | BH790465 | BH790465 SALK 0571 |
| 36 | 6 | 100.0 | 23 | 17 | BH811030 | BH811030 SALK 0571 |
| 37 | 6 | 100.0 | 23 | 17 | BH811032 | BH811032 SALK 0571 |
| 38 | 6 | 100.0 | 24 | 17 | AZ317925 | AZ317925 1M0036G20 |
| 39 | 6 | 100.0 | 24 | 17 | AZ510127 | AZ510127 1M0354L21 |
| 40 | 6 | 100.0 | 24 | 17 | AZ585771 | AZ585771 1M0391L12 |
| 41 | 6 | 100.0 | 24 | 17 | AZ624429 | AZ624429 1M0463D10 |
| 42 | 6 | 100.0 | 24 | 17 | AZ763574 | AZ763574 1M0559P01 |
| 43 | 6 | 100.0 | 24 | 17 | AZ784064 | AZ784064 2M0026N20 |
| 44 | 6 | 100.0 | 24 | 17 | BH846515 | BH846515 SALK 0085 |
| 45 | 6 | 100.0 | 24 | 17 | TA75C02Q | AL458596 T. brucei |

ALIGNMENTS

RESULT 1
AA885444/c
LOCUS
DEFINITION
19 bp mRNA
aml14h07.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1466845 3' similar to TR:00409 000409 CHECKPOINT SUPPRESSOR
1..: mRNA sequence.

ACCESSION
AA885444
VERSION
AA885444.1 GI:2994521
KEYWORDS
EST.
SOURCE
Homo sapiens
ORGANISM
human.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through INM; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert length: 489 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..19

FEATURES
source

```

BASE COUNT
ORIGIN
2 a 8 c 5 g 4 t
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1466845"
/clone_lib="Scares_NFL_T_GBC_S1"
/lab_host="DH10B"
/notes="Organ: pooled; Vector: pTY73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHR, and B-cell
NCL-GCAP-GCBI) were mixed, and 5g circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The diver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731359. Subtraction by Bento
Scares and M. Fatima Bernaldo."

```

| | | | | |
|-----------------------|--------|--------------|----------|---------------|
| Query Match | 100.0% | Score 6: | DB 9: | Length 19: |
| Best Local Similarity | 83.3% | Pred. No. | 4.7e+05: | |
| Matches | 5: | Conservative | 1: | Mismatches 0: |
| | | | | Indels 0: |
| | | | | Gaps 0: |
| Oy | 1 | ATGACU | 6 | |
| | | : | | |
| Db | 12 | ATGACT | 7 | |

| | | | | |
|------------|---|-------|------|------------------------|
| RESULT 2 | | | | |
| BM395830 | | | | |
| LOCUS | BM395830 | 19 bp | mRNA | linear EST 17-JAN-2002 |
| DEFINITION | 5009-0-12-E03.t.1 Chlcoat/Turkewitz cDNA (large fraction) | | | |
| | Tetrahymena thermophila cDNA, mRNA sequence. | | | |

| | | |
|----------|--------------------------|-------------|
| VERSION | BM395830.1 | GI:18195883 |
| KEYWORDS | Est. | |
| SOURCE | Tetrahymena thermophila. | |
| ORGANISM | Tetrahymena thermophila | |

| | |
|-----------|--|
| REFERENCE | 1 (bases 1 to 19) |
| AUTHORS | Turkewitz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel,J., and Klobutcher,L. |
| TITLE | EST from <i>Tetrahymena thermophila</i> , strain CU428.1, growing cells |
| JOURNAL | Unpublished (2002) |
| COMMENT | Contact: Turkewitz AP |

Molecular Genetics and Cell Biology
University of Chicago
920 E. 58th Street, Chicago, IL 60637, USA
Tel: 773 702 4374
Fax: 773 702 3172
Email: apurkewmidway.uchicago.edu
Seq primer: T3.

| FEATURES | source | location/Qualifiers |
|--|--------|---|
| 1..19 | | |
| /organism="Tetranychyna | | thermophila" |
| /strain="CU428.1" | | |
| /db_xref="taxon:5911" | | |
| /clone_lib="Chlicoat/Turkewitz | | cdna (large fraction)" |
| /note="Vector: Bluescript2 SK+; details on library | | preparation can be found in Chlicoat and Turkewitz (2001) |
| Proc. Natl. Acad. Sci USA, 98: 8709-8713.." | | |
| 5 a | 5 c | 6 g 3 t |

```

Query Match      100.0%: Score 6; DB 13; Length 19;
Best Local Similarity 83.3%: Pred. No. 4.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATGACU 6
        |||||:
Db       3 ATGACT 8

```

| | |
|------------|--|
| RESULT | 3 |
| LOCUS | BM396331 |
| DEFINITION | 19 bp mRNA linear EST 17-JAN-2002 5009-0-2-E02.t.1 Chilcoat/Turkewitz cDNA (large fraction) |
| ACCESSION | Tetrahymena thermophila cDNA, mRNA sequence. |
| VERSION | BM396331 |
| KEYWORDS | BM396331.1 GI:18196384 |
| SOURCE | EST. |
| ORGANISM | Tetrahymena thermophila. Tetrahymena thermophila Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida; Tetrahymenina; Tetrahymena. 1 (bases 1 to 19) Turkewitz,A.P., Karrer,K.M., Jahn,C., Ortas,E., Kirk,K.E., Frankel ,J. and Klobutcher,L. EST from Tetrahymena thermophila, strain CU428.1, growing cells Unpublished (2002) |
| REFERENCE | Contact: Turkewitz AP |
| JOURNAL | Molecular Genetics and Cell Biology |
| COMMENT | University Of Chicago 920 E. 58th Street, Chicago, IL 60637, USA Tel: 773 702 4374 Fax: 773 702 3172 Email: apturkew@midway.uchicago.edu Seq primer: T3. |

```

FEATURES
    source
        location/Qualifiers
            1. 19
                /organism="Tetrabymena thermophila"
                /strain="CU428.1"
                /db_xref="taxon:5911"
                /clone_11p="Chilcoat/Turkewitz cDNA (large fraction)"
                /note="Vector: Bluescript2 SK+; details on library
                    preparation can be found in Chilcoat and Turkewitz (2001)
                    Proc. Natl. Acad. Sci USA, 98: 8709-8713."
BASE COUNT
    4 a      6 c      6 g      3 t
ORIGIN

```

| | | | | |
|-----------------------|--------|--------------------|--------|---------------|
| Query Match | 100.0% | Score 6; | DB 13; | Length 19; |
| Best Local Similarity | 83.3% | Pred. No. 4.7e+05; | | |
| Matches | 5; | Conservative | 1; | Mismatches 0; |
| | | | | Indels 0; |
| | | | | Gaps 0; |
| QY | 1 | ATGACU | 6 | |
| | | | | |
| Db | 2 | ATGACT | 7 | |

| LOCUS | DEFINITION | 19 bp | mRNA | linear | EST 17-JAN-2002 |
|----------|--|-------|------|--------|-----------------|
| BM399859 | 5009-b-62-E04.t.2 Chilocoa/Turkewitz cDNA (large fraction) | | | | |
| | tetradymena thermophila cDNA, mRNA sequence. | | | | |

| | |
|-----------|--|
| REFERENCE | |
| ORGANISM | |
| SOURCE | |
| KEYWORDS | |
| VERSION | BM399859 |
| ACCESSION | GI:18199912 |
| | BM399859.1 |
| | EST. |
| | Tetrahymena thermophila. |
| | Tetrahymena thermophila |
| | Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; |
| | Hymenostomatida; Tetrahymenina; Tetrahymena. |

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

I (pages 1 to 19)
Turkewitz A.P., Karrer, K.M., Jahn, C., Ortas, E., Kirk, K.E., Frankel
J., and Klobutcher, L.
EST from Tetrahymena thermophila, strain CU428.1, growing cells
Unpublished (2002)
Contact: Turkewitz AP
Molecular Genetics and Cell Biology
University of Chicago

University of Chicago
920 E. 58th Street, Chicago, IL 60637, USA
Tel: 773 702 4374
Fax: 773 702 3172
Email: apturkew@midway.uchicago.edu
Seq primer: T3.

FEATURES

Location/Qualifiers
1. 19
/organism="Tetrahymena thermophila"
/strain="CU428.1"
/db_xref="taxon:5911"
/clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
/note="Vector: Bluescript SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT
ORIGIN
4 a 4 c 7 g 4 t

Query Match
Best Local Similarity 83.3%; Pred. No. 4.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
11111:
Db 2 ATGACT 7

RESULT 5
D44776 19 bp mRNA linear EST 20-FEB-1998
LOCUS HUMSUP214 Human brain cDNA Homo sapiens CDNA ME51-S-2, mRNA
DEFINITION sequence.
ACCESSION D44776
VERSION D44776.1 GI:1572251
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 19)
Hadano, S., Ishida, Y., Tomiyasu, H., Yamamoto, K., Bates, G.P. and Ikeda, J.
Transcript map of the human chromosome 4p16.3 consisting of 627
CDNA clones derived from 1 Mb of the Huntington's disease locus
DNA Res. 3 (4), 239-255 (1996)
97101646
Contact: Shnji Hadano
Japan Science and Technology Corporation, Neurogenes Project, ICRP
Univ. of Tokai School of Med.
Bohseldal, Isehara, Kanagawa 259-1193, Japan
Tel: 81-463-91-5095
Fax: 81-463-91-4993
Email: shnji@eng.med.u-tokai.ac.jp.

FEATURES
source
1. 19
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MF51-S-2"
/clone_lib="Human brain cDNA"
/issue_type="brain"
/note="Vector: PSORT1; Site: 1: Mui; Site: 2: NotI; mRNA was prepared using a 15mer oligo dt anchored by two
degenerated bases at its 3' end and containing a NotI site
at its 5' end. The cDNA was cloned between Sall and NotI
sites of PSORT1. The Mui-Sall fragment come from the
adaptor used for the cloning. The 3' end is at the NotI
site. cDNA corresponding to abundant species were
eliminated from this library."

BASE COUNT
ORIGIN
5 a 2 c 7 g 5 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 14; Length 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
11111:
Db 13 ATGACT 18

RESULT 6
A2317020/c 19 bp DNA linear GSS 29-SEP-2000
LOCUS IM0035P18F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
DEFINITION clone UUGCIM0035P18 F, DNA sequence.
ACCESSION A2317020
VERSION A2317020.1 GI:10365406
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0035 row: P column: 18
Seq primer: CCGTGTAAACGACGCGCACT
Class: Plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCIM0035P18"
/clone_lib="Mouse 10kb plasmid UUGCIM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42uv. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1473214|g51AR129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
3 a 5 c 4 g 7 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 17; Length 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
11111:
Db 11 ATGACT 6

RESULT 7
 A2457710
 LOCUS 19 bp DNA linear GSS 04-OCT-2000
 DEFINITION IM0261K07F Mouse 10kb plasmid U0GC1M library Mus musculus genomic
 accession A2457710
 clone U0GC1M0261K07 F, DNA sequence.
 VERSION A2457710.1 GI:10615835.
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 19)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 'M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0261 row: K column: 07
 Seq primer: CGTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.
 Location/Qualifiers
 1. 19
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U0GC1M0261K07"
 /clone_lib="Mouse 10kb plasmid U0GC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (g114732114[gb|AF129072.1], a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 8 a 4 c 3 g 4 t
 ORIGIN
 Query Match 100.0%; Score 6; DB 17; Length 19;
 Best Local Similarity 83.3%; Pred. No. 4.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DY 1 ATGACU 6
 |||||
 Db 5 ATGACT 10

RESULT 8
 A2658087
 LOCUS 19 bp DNA linear GSS 14-DEC-2000
 DEFINITION IM0534J09R Mouse 10kb plasmid U0GC1M library Mus musculus genomic
 accession A2658087
 clone U0GC1M0534J09 R, DNA sequence.
 VERSION A2658087.1 GI:11795233
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 19)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 'M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0534 row: J column: 09
 Seq primer: CACACAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 19.
 Location/Qualifiers
 1. 19
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U0GC1M0534J09"
 /clone_lib="Mouse 10kb plasmid U0GC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (g114732114[gb|AF129072.1], a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 7 a 5 c 3 g 4 t
 ORIGIN
 Query Match 100.0%; Score 6; DB 17; Length 19;
 Best Local Similarity 83.3%; Pred. No. 4.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DY 1 ATGACU 6
 |||||
 Db 2 ATGACT 7

RESULT 9
A2663240 19 bp DNA Linear GSS 14-DEC-2000
LOCUS 1M0542118R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0542118 R, DNA sequence.
ACCESSION A2663240
VERSION A2663240.1 GI:11800386
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0542 row: H column: 18
Seq primer: CACACGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0542118"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (q114732114[9b]AF129072.1), a copy number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 7 a 4 c 4 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. NO. 4.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACG 6
|||||
DB 10 ATGACT 15

RESULT 10
A2763729 19 bp DNA Linear GSS 16-FEB-2001
LOCUS 1M0559N16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0559N16 F, DNA sequence.
ACCESSION A2763729
VERSION A2763729.1 GI:12875056
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0559 row: N column: 16
Seq primer: CGTCTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0559N16"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (q114732114[9b]AF129072.1), a copy number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 7 a 3 c 3 g 6 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. NO. 4.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACG 6
|||||
DB 9 ATGACT 14

RESULT 11
 AZ861634
 LOCUS 19 bp DNA linear GSS 21-FEB-2001
 DEFINITION 2M016B18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 accession A2861634
 clone UUGC2M016B18 F, DNA sequence.
 VERSION A2861634
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 19)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellily
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weis,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah
 Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0168 row: B column: 18
 Seq primer: CGTTGTAAACGACGCGCCACT
 Class: plasmid ends
 High quality sequence stop: 19.

FEATURES

source

1. 19
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M016B18"
 /clone_11b="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g114732114|bp|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 BASE COUNT 5 a 4 c 5 g 5 t
 ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
 Best Local Similarity 83.3%; Pred. No. 4.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACU 6
 |||||
 Db 8 ATGACT 13

RESULT 12
 AZ990851
 LOCUS 19 bp DNA linear GSS 27-APR-2001
 DEFINITION 2M0274E15R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
 accession A2990851
 clone UUGC2M0274E15 R, DNA sequence.
 VERSION A2990851
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 19)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellily
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weis,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah
 Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0274 row: E column: 15
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 19.

FEATURES

source

1. 19
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0274E15"
 /clone_11b="Mouse 10kb plasmid UUGC2M library"
 /sex="Female"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (female) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g114732114|bp|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 BASE COUNT 3 a 7 c 4 g 5 t
 ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
 Best Local Similarity 83.3%; Pred. No. 4.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACU 6
 |||||
 Db 7 ATGACT 12

RESULT 13
AZ500712/c 20 bp DNA linear GSS 05-OCT-2000
LOCUS
DEFINITION
1M0339B13F Mouse 10kb plasmid U0GC1M library Mus musculus genomic
clone U0GC1M0339B13 F, DNA sequence.
ACCESSION
AZ500712
VERSION
AZ500712.1 GI:10680801
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
TITLE
Unpublished (2000)
JOURNAL
Contact: Robert B. Weiss
COMMENT
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0339 row: B column: 13
Seq primer: CGTTGTAACGACGCCACG
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1..20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U0GC1M0339B13"
/clone_1lb="Mouse 10kb plasmid U0GC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD2 (gii4732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 7 a 6 c 1 g 6 t
ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 4.9e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 11 ATGACT 6

RESULT 14
AZ830285 20 bp DNA linear GSS 20-FEB-2001
LOCUS
DEFINITION
2M0109N22F Mouse 10kb plasmid U0GC1M library Mus musculus genomic
clone U0GC2M0109N22 F, DNA sequence.
ACCESSION
AZ830285
VERSION
AZ830285.1 GI:13000193
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
TITLE
Unpublished (2000)
JOURNAL
Contact: Robert B. Weiss
COMMENT
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0109 row: N column: 22
Seq primer: CGTTGTAACGACGCCACG
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1..20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U0GC2M0109N22"
/clone_1lb="Mouse 10kb plasmid U0GC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD2 (gii4732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 4 a 6 c 1 g 4 t
ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 4.9e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 11 ATGACT 6

RESULT 15

AZ455887

LOCUS

21 bp DNA linear GSS 04-OCT-2000

DEFINITION

1M02358019F Mouse 10kb plasmid U0GCM1 library Mus musculus genomic

ACCESSION

AZ455887

clone U0GCM0258019 F, DNA sequence.

VERSION

AZ455887

GSS.

KEYWORDS

house mouse.

REFERENCE

1 (bases 1 to 21)

Dunn, P., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

AUTHORS

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE

plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0258 row: 0 column: 19

Seq primer: CGTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 21.

FEATURES

Location/Qualifiers

1..21

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="U0GCM0258019"

/clone_lib="Mouse 10kb plasmid U0GCM1 library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (g1473214[gb|AF129072.1]), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT

6 a 4 c 4 g 7 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 21;
Best Local Similarity 83.3%; Pred. No. 5e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6

|||||:

Db 3 ATGACT 8

Search completed: May 21, 2003, 06:25:59
Job time : 1100 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 05:13:24 ; Search time 331.5 Seconds

(without alignments)
526.748 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| C 1 | 6 | 100.0 | 10 | 6 | AR128987 | AR128987 Sequence |
| C 2 | 6 | 100.0 | 10 | 6 | AR128988 | AR128988 Sequence |
| C 3 | 6 | 100.0 | 10 | 6 | AR128989 | AR128989 Sequence |
| C 4 | 6 | 100.0 | 10 | 6 | AR128990 | AR128990 Sequence |
| C 5 | 6 | 100.0 | 10 | 6 | AR128991 | AR128991 Sequence |
| C 6 | 6 | 100.0 | 10 | 6 | AX152349 | AX152349 Sequence |
| C 7 | 6 | 100.0 | 10 | 6 | AX152410 | AX152410 Sequence |
| C 8 | 6 | 100.0 | 10 | 6 | AX152849 | AX152849 Sequence |
| C 9 | 6 | 100.0 | 10 | 6 | AX152886 | AX152886 Sequence |
| C 10 | 6 | 100.0 | 10 | 6 | AX153384 | AX153384 Sequence |
| C 11 | 6 | 100.0 | 10 | 6 | AX153513 | AX153513 Sequence |
| C 12 | 6 | 100.0 | 10 | 6 | E16890 | E16890 DNA sequence |
| C 13 | 6 | 100.0 | 11 | 6 | A91501 | A91501 Sequence 28 |
| C 14 | 6 | 100.0 | 11 | 6 | A91503 | A91503 Sequence 30 |
| C 15 | 6 | 100.0 | 11 | 6 | AX470495 | AX470495 Sequence |
| C 16 | 6 | 100.0 | 11 | 6 | AX470740 | AX470740 Sequence |
| C 17 | 6 | 100.0 | 11 | 6 | AX470747 | AX470747 Sequence |
| C 18 | 6 | 100.0 | 11 | 6 | AX471502 | AX471502 Sequence |
| C 19 | 6 | 100.0 | 11 | 6 | AX471630 | AX471630 Sequence |
| C 20 | 6 | 100.0 | 11 | 6 | AX471677 | AX471677 Sequence |
| C 21 | 6 | 100.0 | 11 | 6 | AX471805 | AX471805 Sequence |
| C 22 | 6 | 100.0 | 12 | 6 | A06058 | A06058 Synthetic P |
| C 23 | 6 | 100.0 | 12 | 6 | A06059 | A06059 Synthetic P |
| C 24 | 6 | 100.0 | 12 | 6 | A15123 | A15123 Nucleotide |
| C 25 | 6 | 100.0 | 12 | 6 | A16601 | A16601 Nucleotide |
| C 26 | 6 | 100.0 | 12 | 6 | A16602 | A16602 Nucleotide |
| C 27 | 6 | 100.0 | 12 | 6 | A47656 | A47656 Sequence 16 |
| C 28 | 6 | 100.0 | 12 | 6 | A61520 | A61520 Sequence 89 |
| C 29 | 6 | 100.0 | 12 | 6 | A91489 | A91489 Sequence 16 |
| C 30 | 6 | 100.0 | 12 | 6 | AR027874 | AR027874 Sequence |
| C 31 | 6 | 100.0 | 12 | 6 | AR075465 | AR075465 Sequence |
| C 32 | 6 | 100.0 | 12 | 6 | AR075468 | AR075468 Sequence |
| C 33 | 6 | 100.0 | 12 | 6 | AR101000 | AR101000 Sequence |
| C 34 | 6 | 100.0 | 12 | 6 | AR137925 | AR137925 Sequence |
| C 35 | 6 | 100.0 | 12 | 6 | AR153916 | AR153916 Sequence |
| C 36 | 6 | 100.0 | 12 | 6 | AR153919 | AR153919 Sequence |
| C 37 | 6 | 100.0 | 12 | 6 | AR167701 | AR167701 Sequence |
| C 38 | 6 | 100.0 | 12 | 6 | AR178533 | AR178533 Sequence |
| C 39 | 6 | 100.0 | 12 | 6 | AR178536 | AR178536 Sequence |
| C 40 | 6 | 100.0 | 12 | 6 | AR199100 | AR199100 Sequence |
| C 41 | 6 | 100.0 | 12 | 6 | AX233626 | AX233626 Sequence |
| C 42 | 6 | 100.0 | 12 | 6 | E29585 | E29585 Method for |
| C 43 | 6 | 100.0 | 12 | 6 | E38691 | E38691 Method for |
| C 44 | 6 | 100.0 | 12 | 6 | E64117 | E64117 Method for |
| C 45 | 6 | 100.0 | 12 | 6 | I07918 | I07918 Sequence 30 |

ALIGNMENTS

RESULT 1
AR128987/c
LOCUS AR128987 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6183966.
ACCESSION AR128987
VERSION AR128987.1 GI:14116649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLES Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 2 06-FEB-2001.

FEATURES location/Qualifiers
source 1. .10
/organism="unknown"
BASE COUNT 2 a 3 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 6 GCCTCT 1

RESULT 2
LOCUS AR128988 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6183966.
ACCESSION AR128988
VERSION AR128988.1 GI:14116650
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 3 06-FEB-2001;
FEATURES Location/Qualifiers
source 1. .10
/organism="unknown"

BASE COUNT 2 a 3 c 5 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 7 GCCTCT 2

RESULT 3
LOCUS AR128989 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 4 from patent US 6183966.
ACCESSION AR128989
VERSION AR128989.1 GI:14116651
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 4 06-FEB-2001;
FEATURES Location/Qualifiers
source 1. .10
/organism="unknown"

BASE COUNT 3 a 3 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 4
LOCUS AR128990 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6183966.
ACCESSION AR128990
VERSION AR128990.1 GI:14116652
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 5 06-FEB-2001;
FEATURES Location/Qualifiers
source 1. .10
/organism="unknown"

BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 9 GCCTCT 4

RESULT 5
LOCUS AR128991 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6183966.
ACCESSION AR128991
VERSION AR128991.1 GI:14116653
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 6 06-FEB-2001;
FEATURES Location/Qualifiers
source 1. .10
/organism="unknown"

BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 10 GCCTCT 5

RESULT 6
LOCUS AX152349 10 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 264 from Patent W00138577.
ACCESSION AX152349
VERSION AX152349.1 GI:14534000
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 264 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source
1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 0 a 5 c 2 g 3 t
ORIGIN

Query Match
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 1 GCCTCT 6

RESULT 7
AX152410 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 325 from Patent WO0138577.
DEFINITION AX152410
ACCESSION AX152410
VERSION AX152410.1 GI:14534061
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 325 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source
1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1 a 5 c 2 g 2 t
ORIGIN

Query Match
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 1 GCCTCT 6

RESULT 8
AX152849 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 764 from Patent WO0138577.
DEFINITION AX152849
ACCESSION AX152849
VERSION AX152849.1 GI:14534500
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 764 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source
1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1 a 3 c 3 g 3 t
ORIGIN

Query Match
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 5 GCCTCT 10

RESULT 9
AX152886/c 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 801 from Patent WO0138577.
DEFINITION AX152886
ACCESSION AX152886
VERSION AX152886.1 GI:14534537
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 801 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source
1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 5 a 1 c 3 g 1 t
ORIGIN

Query Match
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 10
AX153384 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 1299 from Patent WO0138577.
DEFINITION AX153384
ACCESSION AX153384
VERSION AX153384.1 GI:14535035
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 1299 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source
1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN

Query Match
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 1 GCCTCT 6

Db 2 GCCTCT 7

RESULT 11
AX153513
LOCUS AX153513 10 bp DNA
DEFINITION Sequence 1428 from Patent WO0138577.
ACCESSION AX153513
VERSION AX153513.1 GI:14535164
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
Velculescu, V.E., Vogelstein, B. and Kinzler, K.W.
Human transcriptomes
Patent: WO 0138577-A 1428 31-MAY-2001;
The Johns Hopkins University (US)
Location/Qualifiers
1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 0 a 4 c 2 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||

Db 1 GCCTCT 6

RESULT 12
E16890
LOCUS E16890 10 bp DNA
DEFINITION DNA sequence required for efficient protein transcription in
Brevibacterium flavum.
E16890
ACCESSION E16890.1 GI:5711573
VERSION E16890.1
KEYWORDS JP 1998229881-A/31.
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1 (bases 1 to 10)
Kobayashi, M., Man, T. and Yugawa, H.
DNA HAVING SEQUENCE CAPABLE OF EFFICIENTLY TRANSLATING PROTEIN IN
CORYNEFORM BACTERIA
Patent: JP 1998229881-A 31 02-SEP-1998;
MITSUBISHI CHEM CORP
OS Brevibacterium flavum
PD JP 1998229881-A/31
PN 02-SEP-1998
PI 19-FEB-1997 JP 1997035338
PT KOBAYASHI MIKI, MAN TOMOKO, YUGAWA HIDEAKI
PC C12N15/09, C07H21/04, C12N1/21//C12N9/38, C12Q1/68, (C12N15/09, PC
C12N1/19)
PC (C12N1/21, C12N1/13), (C12N9/38, C12N1/19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1..10
/strain="MJ-233",
Location/Qualifiers
1..10
/organism="Corynebacterium glutamicum"

FEATURES
source

BASE COUNT 3 a 1 c 5 g 1 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||

Db 9 GCCTCT 4

RESULT 13
A91501
LOCUS A91501 11 bp DNA
DEFINITION Sequence 28 from Patent WO9824928.
ACCESSION A91501
VERSION A91501.1 GI:6740456
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 11)
AUTHORS Pallsgaard, N. and Hokland, P.
TITLE DETECTION OF CHROMOSOMAL ABNORMALITIES
JOURNAL Patent: WO 9824928-A 28 11-JUN-1998;
PALLSGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
1..11
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 1 a 5 c 2 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2.7e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||

Db 5 GCCTCT 10

RESULT 14
A91503
LOCUS A91503 11 bp DNA
DEFINITION Sequence 30 from Patent WO9824928.
ACCESSION A91503
VERSION A91503.1 GI:6740458
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 11)
AUTHORS Pallsgaard, N. and Hokland, P.
TITLE DETECTION OF CHROMOSOMAL ABNORMALITIES
JOURNAL Patent: WO 9824928-A 30 11-JUN-1998;
PALLSGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
1..11
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 0 a 5 c 2 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2.7e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||

Db 5 GCCTCT 10

RESULT 15
AX470495/c 11 bp DNA linear PAT 09-AUG-2002
LOCUS AX470495
DEFINITION Sequence 72 from Patent WO02053773.
ACCESSION AX470495
VERSION AX470495.1 GI:22205620
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Hofmann, K., Conradt, M. and Petersohn, D.
TITLE Method for determining skin stress or skin ageing in vitro
JOURNAL Patent: WO 02053773-A 72 11-JUL-2002;
HENKEL KGAA (DE)
FEATURES
source location/Qualifiers
1. 11
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 6 a 2 c 3 g 0 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. NO. 2.7e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3
Search completed: May 21, 2003, 07:17:18
Job time : 332.5 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:50:53 ; Search time 148.25 Seconds
(without alignments)
91.143 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table: OLIGO_NUC

Searched: 2185239 seqs, 112599159 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N.GeneSeq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 6 | 100.0 | 8 | 20 | AAx29578 |
| 2 | 6 | 100.0 | 8 | 21 | AAx80736 |
| 3 | 6 | 100.0 | 9 | 24 | ABO71469 |
| 4 | 6 | 100.0 | 9 | 24 | ABO71786 |
| 5 | 6 | 100.0 | 9 | 24 | ABO71788 |
| 6 | 6 | 100.0 | 9 | 24 | ABO71921 |
| 7 | 6 | 100.0 | 9 | 24 | ABO71982 |
| 8 | 6 | 100.0 | 9 | 24 | ABO72173 |
| 9 | 6 | 100.0 | 9 | 24 | ABO72175 |

| | | | | | | |
|------|---|-------|----|----|----------|--------------------|
| c 10 | 6 | 100.0 | 10 | 16 | AAQ96863 | HIV-1 NL4-3 nef ge |
| c 11 | 6 | 100.0 | 10 | 16 | AAQ96864 | HIV-1 NL4-3 nef ge |
| c 12 | 6 | 100.0 | 10 | 16 | AAQ96865 | HIV-1 NL4-3 nef ge |
| c 13 | 6 | 100.0 | 10 | 16 | AAQ96866 | HIV-1 NL4-3 nef ge |
| c 14 | 6 | 100.0 | 10 | 16 | AAQ96867 | HIV-1 NL4-3 nef ge |
| c 15 | 6 | 100.0 | 10 | 20 | AAx86201 | SAGE tag used to 1 |
| c 16 | 6 | 100.0 | 10 | 20 | AAx00518 | Antisense oligonuc |
| c 17 | 6 | 100.0 | 10 | 20 | AAx00519 | Antisense oligonuc |
| c 18 | 6 | 100.0 | 10 | 20 | AAx00520 | Antisense oligonuc |
| c 19 | 6 | 100.0 | 10 | 20 | AAx00520 | Antisense oligonuc |
| c 20 | 6 | 100.0 | 10 | 20 | AAx00521 | Antisense oligonuc |
| c 21 | 6 | 100.0 | 10 | 21 | AAx56532 | Human macrophage g |
| c 22 | 6 | 100.0 | 10 | 21 | AAx77725 | Human dendritic ce |
| c 23 | 6 | 100.0 | 10 | 21 | AAx77746 | Human dendritic ce |
| c 24 | 6 | 100.0 | 10 | 21 | AAx77927 | Human dendritic ce |
| c 25 | 6 | 100.0 | 10 | 21 | AAx78138 | Human dendritic ce |
| c 26 | 6 | 100.0 | 10 | 21 | AAx78540 | Human dendritic ce |
| c 27 | 6 | 100.0 | 10 | 21 | AAx78579 | Human dendritic ce |
| c 28 | 6 | 100.0 | 10 | 21 | AAx78692 | Human dendritic ce |
| c 29 | 6 | 100.0 | 10 | 21 | AAx78858 | Human dendritic ce |
| c 30 | 6 | 100.0 | 10 | 21 | AAx79181 | Human dendritic ce |
| c 31 | 6 | 100.0 | 10 | 21 | AAx79724 | Human colon tumour |
| c 32 | 6 | 100.0 | 10 | 21 | AAx79825 | Human lung prefe |
| c 33 | 6 | 100.0 | 10 | 21 | AAx79884 | Human dendritic ce |
| c 34 | 6 | 100.0 | 10 | 21 | AAx81089 | Metastatic breast |
| c 35 | 6 | 100.0 | 10 | 21 | AAx81276 | Metastatic breast |
| c 36 | 6 | 100.0 | 10 | 21 | AAx81485 | Metastatic breast |
| c 37 | 6 | 100.0 | 10 | 21 | AAx81751 | Metastatic breast |
| c 38 | 6 | 100.0 | 10 | 21 | AAx81798 | Metastatic breast |
| c 39 | 6 | 100.0 | 10 | 21 | AAx82217 | Metastatic breast |
| c 40 | 6 | 100.0 | 10 | 21 | AAx82317 | Metastatic breast |
| c 41 | 6 | 100.0 | 10 | 21 | AAx82473 | Metastatic breast |
| c 42 | 6 | 100.0 | 10 | 21 | AAx82783 | Metastatic breast |
| c 43 | 6 | 100.0 | 10 | 21 | AAx82938 | Metastatic breast |
| c 44 | 6 | 100.0 | 10 | 21 | AAx83462 | Metastatic breast |
| c 45 | 6 | 100.0 | 10 | 21 | AAx84090 | Metastatic breast |

ALIGNMENTS

| | | |
|----------|---|----------------------|
| RESULT 1 | AAx29578 | standard; DNA; 8 BP. |
| ID | AAx29578 | |
| XX | AAx29578; | |
| AC | 03-JUN-1999 (first entry) | |
| XX | | |
| XX | | |
| DT | Primer for human G-protein coupled receptor genes. | |
| DE | | |
| XX | | |
| XX | Nucleic acid amplification; nuclear receptor; G-protein coupled receptor; | |
| KW | apoptosis; DNA repair; DNA replication; plant biology; agriculture; | |
| KW | human; veterinary medicine; reproduction; microbiology; hybridisation; | |
| KW | environmental science; DNA fingerprinting; PCR primer; ss. | |
| OS | Synthetic. | |
| OS | Homo sapiens. | |
| XX | | |
| XX | | |
| PN | W09911823-A2. | |
| XX | | |
| PD | 11-MAR-1999. | |
| XX | | |
| PF | 04-SEP-1998; 98WO-US18392. | |
| XX | | |
| PR | 05-SEP-1997; 97US-0925816. | |
| XX | | |
| PA | (KIMM-) KIMMEL CANCER CENT SIDNEY. | |
| XX | | |
| PI | McClelland M, Pesole G; | |
| XX | | |
| DR | WPI; 1999-205200/17. | |
| XX | | |

PT Subset of primers able to amplify group of related sequences
 XX
 PS Claim 21; Page 75; 92pp; English.
 XX
 CC The invention provides primers (AA29501-X29679) for identifying
 CC sequences encoding structurally or functionally related proteins such as
 CC nuclear or G-protein coupled receptors, apoptosis-related or DNA
 CC repair/replication proteins. The identified sequences are broadly useful
 CC in plant biology, agriculture, human or veterinary medicine,
 CC reproduction, microbiology or environmental science, e.g. to study
 CC expression of nuclear receptors at different stages of tissue development
 CC or after treatment with particular drugs. It is also used for DNA
 CC fingerprinting (to generate products useful for differential
 CC hybridisation), or, where a 3'-anchor primer is used, to isolate the
 CC 3'-ends of mRNA sequences. Sequences AA29526-X29583 represent claimed
 CC primers specific for human G-protein coupled receptor genes.
 CC
 SQ Sequence 8 BP; 0 A; 4 C; 2 G; 2 T; 0 other;
 Query Match 100.0%; Score 6; DB 20; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.6e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCT 6
 Db 1 GCCTCT 6
 RESULT 2
 AAA80736
 ID AAA80736 standard; DNA: 8 BP.
 AC AAA80736;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE A. thaliana primer walking octamer SEQ ID NO: 49.
 XX
 KW Primer walking; octamer; primer; DNA sequencing; PCR; ss.
 OS
 XX Arabidopsis thaliana.
 OS
 XX US6083695-A.
 PN
 XX 04-JUL-2000.
 PD
 XX 21-MAY-1997; 97US-0859954.
 PF
 XX 15-APR-1996; 96US-0632782.
 PR
 XX
 PA (UYHO-) UNIV HOUSTON.
 XX
 PA (HARD/) HARDIN S H.
 XX
 PI Hardin PE, Hardin SH, Homayouni R;
 XX
 DR WPI; 2000-474852/41.
 XX
 PT Sequencing an unknown DNA molecule for the polymerase chain reaction
 PT and other primer processes comprises primer walking of octamer
 PT oligonucleotides -
 PS
 XX Example 8; Column 49-50; 161pp; English.
 PS
 CC This invention describes a novel method for sequencing an unknown DNA
 CC molecule which comprises selecting a library primer from an octamer
 CC oligonucleotide library consisting of 48 8-bp sequences and
 CC corresponding complementary sequences, where the library primer is
 CC complementary to a known sequence adjacent to the unknown sequence or
 CC is complementary to a sequence in a known extension product. The method
 CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
 CC which make use of primers. The octamers are used to identify coding
 CC sequences. Primer walking using the octamer libraries is advantageous
 CC over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
 CC directed and methodical approach. AA60688-A61253 represent the octamer
 CC primers used in the primer walking method of the invention.
 CC
 SQ Sequence 8 BP; 0 A; 3 C; 1 G; 4 T; 0 other;
 Query Match 100.0%; Score 6; DB 21; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.6e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCT 6
 Db 3 GCCTCT 8
 RESULT 3
 ABQ71469/C
 ID ABQ71469 standard; DNA: 9 BP.
 AC ABQ71469;
 XX
 DT 28-AUG-2002 (first entry)
 XX
 DE Zinc finger protein related oligonucleotide target SEQ ID NO:588.
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 XX WO200242459-A2.
 PN
 XX 30-MAY-2002.
 PD
 XX 20-NOV-2001; 2001WO-US43438.
 PF
 XX 20-NOV-2000; 2000US-0716637.
 PR
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA
 XX Liu Q;
 PI
 XX WPI; 2002-500284/53.
 DR
 XX
 PT New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 PS
 XX Example 1; Page 44; 81pp; English.
 PS
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP46191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;
SQ

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.3e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 4
ABQ71786/c
ID ABQ71786 standard; DNA; 9 BP.
XX
AC ABQ71786;
XX
DT 28-AUG-2002 (first entry)
XX

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2084.
XX
KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200242459-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Liu Q;
XX
DR WPI; 2002-500284/53.
XX

PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
PS Example 1; Page 55; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
CC (II) comprising (1); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP1230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.3e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 5
ABQ71788/c
ID ABQ71788 standard; DNA; 9 BP.
XX
AC ABQ71788;
XX
DT 28-AUG-2002 (first entry)
XX

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2086.
XX
KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200242459-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Liu Q;
XX
DR WPI; 2002-500284/53.
XX

PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
PS Example 1; Page 55; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
CC (II) comprising (1); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP1230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.3e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 6

ABQ71921/c
ID ABQ71921 standard; DNA; 9 BP.

AC ABQ71921;

DT 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2219.

KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN W0200242459-A2.

PD 30-MAY-2002.

PF 20-NOV-2001; 2001WO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying gene function and for human therapeutics and plant engineering, comprises first, second and third zinc fingers, ordered from N- to C-terminus

PS Example 1; Page 58; 81bp; English.

CC The present invention describes a zinc finger protein (I) that binds to a target site, comprising a first (F1), a second (F2), and a third (F3) zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the target site comprises, in 3'-5' direction, a first (S1), a second (S2), and a third (S3) target site. Also described are: (1) a polypeptide (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and (3) designing (M) (I) involves selecting the F1 zinc finger such that it binds to the S1 target site, selecting the F2 zinc finger such that it binds to the S2 target site, and selecting the F3 zinc finger such that it binds to the S3 target site, thus designing (I) that binds to a target site. (I) is useful for recognition of triplet target sites having the nucleotide G in the 5'-most position of the target site. (I) is useful in studying gene function, and for human therapeutics and plant engineering. (I), (II) or (III) is useful in therapeutic methods to modulate the expression of a target region within a subject, in diagnostic methods for sequence specific detection of target nucleic acid in a sample, and in assays to determine the phenotype and function of gene expression. (I) has improved affinity and specificity for their target sequences, as well as enhanced biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230 represent DNA target sequences and zinc finger peptides which are given in the exemplification of the present invention.

CC Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.3e+08;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:

Db 8 GCCTCT 3

RESULT 7

ABQ71982/c
ID ABQ71982 standard; DNA; 9 BP.

AC ABQ71982;

DT 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2280.

KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN W0200242459-A2.

PD 30-MAY-2002.

PF 20-NOV-2001; 2001WO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying gene function and for human therapeutics and plant engineering, comprises first, second and third zinc fingers, ordered from N- to C-terminus

PS Example 1; Page 59; 81bp; English.

CC The present invention describes a zinc finger protein (I) that binds to a target site, comprising a first (F1), a second (F2), and a third (F3) zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the target site comprises, in 3'-5' direction, a first (S1), a second (S2), and a third (S3) target site. Also described are: (1) a polypeptide (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and (3) designing (M) (I) involves selecting the F1 zinc finger such that it binds to the S1 target site, selecting the F2 zinc finger such that it binds to the S2 target site, and selecting the F3 zinc finger such that it binds to the S3 target site, thus designing (I) that binds to a target site. (I) is useful for recognition of triplet target sites having the nucleotide G in the 5'-most position of the target site. (I) is useful in studying gene function, and for human therapeutics and plant engineering. (I), (II) or (III) is useful in therapeutic methods to modulate the expression of a target region within a subject, in diagnostic methods for sequence specific detection of target nucleic acid in a sample, and in assays to determine the phenotype and function of gene expression. (I) has improved affinity and specificity for their target sequences, as well as enhanced biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230 represent DNA target sequences and zinc finger peptides which are given in the exemplification of the present invention.

CC Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.3e+08;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

```

RESULT 8
AB072173/c
ID AB072173 standard; DNA: 9 BP.
XX
XX AB072173:
AC
XX 28-AUG-2002 (first entry)
DE
XX
XX Zinc finger protein related oligonucleotide target SEQ ID NO:2471.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX MO200242459-A2.
XX
XX 30-MAY-2002.
XX
XX 20-NOV-2001; 2001MO-US43438.
XX
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Liu Q:
XX
XX WPI; 2002-500284/53.
XX
XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering,
XX comprises first, second and third zinc fingers, ordered from N- to
XX C-terminus
XX
XX Example 1; Page 63; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target subsequence. Also described are: (i) a polypeptide
XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target subsequence, selecting the F2 zinc finger such
XX that it binds to the S2 target subsequence, and selecting the F3 zinc
XX finger such that it binds to the S3 target subsequence, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target subsequences having the nucleotide G in the 5'-most position of the
XX subsequence. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.
XX
XX Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
SQ
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 6; DB 24; Length 9;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCCTCU 6
XX 11111:
XX Db 8 GCCTCT 3
XX
XX RESULT 9
XX AB072175/c
XX ID AB072175 standard; DNA: 9 BP.

```

```

XX
XX AC AB072175:
XX
XX 28-AUG-2002 (first entry)
DE
XX
XX Zinc finger protein related oligonucleotide target SEQ ID NO:2473.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX MO200242459-A2.
XX
XX 30-MAY-2002.
XX
XX 20-NOV-2001; 2001MO-US43438.
XX
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Liu Q:
XX
XX WPI; 2002-500284/53.
XX
XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering,
XX comprises first, second and third zinc fingers, ordered from N- to
XX C-terminus
XX
XX Example 1; Page 63; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target subsequence. Also described are: (i) a polypeptide
XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target subsequence, selecting the F2 zinc finger such
XX that it binds to the S2 target subsequence, and selecting the F3 zinc
XX finger such that it binds to the S3 target subsequence, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target subsequences having the nucleotide G in the 5'-most position of the
XX subsequence. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.
XX
XX Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
SQ
XX
XX Query Match
XX Best Local Similarity 83.3%; Score 6; DB 24; Length 9;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCCTCU 6
XX 11111:
XX Db 8 GCCTCT 3
XX
XX RESULT 10
XX AA096863/c
XX ID AA096863 standard; DNA: 10 BP.
XX
XX AC AA096863:
XX

```

```

DT 26-MAR-1996 (first entry)
XX
XX HIV-1 NL4-3 nef gene nucleotide deletion 458.
DE
XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
KW
XX Human immunodeficiency virus type 1.
OS
XX MO9521912-A1.
PN
XX 17-AUG-1995.
PD
XX 14-FEB-1995; 95WO-AU00063.
PF
XX 23-DEC-1994; 94AU-0000284.
PR 14-FEB-1994; 94AU-0003864.
PR 21-FEB-1994; 94AU-0004002.
XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
PS
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AAQ96406-Q97018) from the nef gene and/or
CC 1 or more decanucleotides (AAQ97019-Q97166) from the LTR region; the
CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
CC (AAQ96141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
XX Sequence 10 BP; 4 A; 1 C; 4 G; 1 T; 0 other;
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
DB 10 GCCTCT 5

```

```

XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
PS
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AAQ96406-Q97018) from the nef gene and/or
CC 1 or more decanucleotides (AAQ97019-Q97166) from the LTR region; the
CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
CC (AAQ96141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
XX Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
DB 9 GCCTCT 4

```

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 83.3%; Pred. No. 2.9e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCT 6
 DB 8 GCCTCT 3

RESULT 13
 AA096866/c
 ID AA096866 standard; DNA; 10 BP.

AC AA096866;

XX 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 461.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX 17-AUG-1995.

PD 14-FEB-1995; 95WO-AU00063.

XX 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

XX 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

DR WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

XX Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deancucleotides (AA096406-097018) from the nef gene and/or

CC 1 or more deancucleotides (AA097019-097166) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene

CC (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

SQ Sequence 10 BP; 5 A; 2 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 83.3%; Pred. No. 2.9e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCT 6

DB 7 GCCTCT 2

RESULT 14

AA096867/c
 ID AA096867 standard; DNA; 10 BP.

XX AA096867;

XX 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 462.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX 17-AUG-1995.

PD 14-FEB-1995; 95WO-AU00063.

XX 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

XX 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

DR WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

XX Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deancucleotides (AA096406-097018) from the nef gene and/or

CC 1 or more deancucleotides (AA097019-097166) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene

CC (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

SQ Sequence 10 BP; 4 A; 2 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 83.3%; Pred. No. 2.9e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCT 6

DB 6 GCCTCT 1

RESULT 15

AA096201

ID AA096201 standard; DNA; 10 BP.

AC AA096201;

XX 22-SEP-1999 (first entry)

DE SAGE tag used to identify transcripts which are enhanced by p53.

XX p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity;

XX neoplastic; p53 binding site; PIG-3 promoter; SAGE tag; ss.

OS Homo sapiens.

XX WO9914356-A2.

PN 25-MAR-1999.

PD

XX 17-SEP-1998; 98WO-US19300.
 PF 30-MAR-1998; 98US-0079817.
 PR 17-SEP-1997; 97US-0059153.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX
 PI Kinzler KW, Polyak K, Vogelstein B;
 XX
 DR WPI; 1999-443793/37.
 XX

PT Use of p53 transcription tags to determine p53 status in, e.g.
 PF cancer diagnosis
 XX

PS Example 1; Page 25; 73pp; English.
 XX

CC The specification describes the use of p53 transcription tags for
 CC developing products to determine p53 status, to diagnose cancer
 CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
 CC A method for diagnosing cancer or determining p53 status in a sample
 CC suspected for being neoplastic comprises comparing the level of
 CC transcription of an RNA transcript in a first sample (s1) of a first
 CC tissue (t1) to the level of transcription of the transcript in a second
 CC sample (s2) of a second tissue (s2), where s1 is suspected of being
 CC neoplastic and s2 is a normal human tissue (of the same type) and the
 CC transcript is identified by a tag; and categorizing s1 as neoplastic
 CC or as having a mutant p53 when transcription is found to be the same
 CC or lower in the first, than in s2. The methods and products can be used
 CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
 CC or carcinogenicity of a test agent. AAX86201-33 represent SAGE tags
 CC used to identify transcripts which are enhanced by p53.
 XX

SQ Sequence 10 BP; 0 A; 6 C; 1 G; 3 T; 0 other;
 XX

Query Match 100.0%; Score 6; DB 20; Length 10;
 Best Local Similarity 83.3%; Pied. NO. 2.9e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
 |||||
 Db 4 GCCTCT 9

Search completed: May 21, 2003, 06:54:55
 Job time : 149.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:28:43 ; Search time 34.75 Seconds
(without alignments)
52.951 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6
Sequence: 1 gcctcu 6

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued_patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCBUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 6 | 100.0 | 8 | 3 | US-08-859-954-49 |
| 2 | 6 | 100.0 | 10 | 3 | US-08-388-353-459 |
| 3 | 6 | 100.0 | 10 | 3 | US-08-388-353-460 |
| 4 | 6 | 100.0 | 10 | 3 | US-08-388-353-461 |
| 5 | 6 | 100.0 | 10 | 3 | US-08-388-353-462 |
| 6 | 6 | 100.0 | 10 | 3 | US-08-388-353-463 |
| 7 | 6 | 100.0 | 10 | 3 | US-08-488-551B-459 |
| 8 | 6 | 100.0 | 10 | 3 | US-08-488-551B-460 |
| 9 | 6 | 100.0 | 10 | 3 | US-08-488-551B-461 |
| 10 | 6 | 100.0 | 10 | 3 | US-08-488-551B-462 |
| 11 | 6 | 100.0 | 10 | 3 | US-08-488-551B-463 |
| 12 | 6 | 100.0 | 10 | 4 | US-09-235-614-2 |
| 13 | 6 | 100.0 | 10 | 4 | US-09-235-614-3 |
| 14 | 6 | 100.0 | 10 | 4 | US-09-235-614-4 |
| 15 | 6 | 100.0 | 10 | 4 | US-09-235-614-5 |
| 16 | 6 | 100.0 | 10 | 4 | US-09-235-614-6 |
| 17 | 6 | 100.0 | 10 | 4 | US-09-154-750A-4 |
| 18 | 6 | 100.0 | 12 | 2 | US-08-494-301A-16 |
| 19 | 6 | 100.0 | 12 | 2 | US-08-480-020B-18 |
| 20 | 6 | 100.0 | 12 | 2 | US-08-480-020B-21 |
| 21 | 6 | 100.0 | 12 | 2 | US-08-910-618-18 |
| 22 | 6 | 100.0 | 12 | 2 | US-08-910-618-21 |
| 23 | 6 | 100.0 | 12 | 3 | US-08-779-355-5 |
| 24 | 6 | 100.0 | 12 | 3 | US-08-671-824-19 |
| 25 | 6 | 100.0 | 12 | 3 | US-08-874-825-88 |
| 26 | 6 | 100.0 | 12 | 3 | US-08-938-835A-5 |
| 27 | 6 | 100.0 | 12 | 3 | US-08-663-824-88 |

| | | | | | | |
|----|---|-------|----|---|--------------------|-------------------|
| 28 | 6 | 100.0 | 12 | 4 | US-09-243-335-1 | Sequence 1, Appl |
| 29 | 6 | 100.0 | 12 | 4 | US-08-910-322-18 | Sequence 18, Appl |
| 30 | 6 | 100.0 | 12 | 4 | US-08-910-322-21 | Sequence 21, Appl |
| 31 | 6 | 100.0 | 12 | 4 | US-09-281-418-65 | Sequence 65, Appl |
| 32 | 6 | 100.0 | 12 | 4 | US-08-484-939A-18 | Sequence 18, Appl |
| 33 | 6 | 100.0 | 12 | 4 | US-08-484-939A-21 | Sequence 21, Appl |
| 34 | 6 | 100.0 | 12 | 4 | US-09-043-149-48 | Sequence 48, Appl |
| 35 | 6 | 100.0 | 12 | 4 | US-09-231-303-88 | Sequence 88, Appl |
| 36 | 6 | 100.0 | 13 | 1 | US-08-284-746-5 | Sequence 5, Appl |
| 37 | 6 | 100.0 | 13 | 1 | US-08-284-746-16 | Sequence 16, Appl |
| 38 | 6 | 100.0 | 13 | 1 | US-08-050-073-162 | Sequence 162, App |
| 39 | 6 | 100.0 | 14 | 1 | US-08-303-004-23 | Sequence 23, Appl |
| 40 | 6 | 100.0 | 14 | 1 | US-08-050-073-99 | Sequence 99, Appl |
| 41 | 6 | 100.0 | 14 | 1 | US-08-050-073-161 | Sequence 161, App |
| 42 | 6 | 100.0 | 14 | 1 | US-08-192-941-25 | Sequence 25, App |
| 43 | 6 | 100.0 | 14 | 1 | US-08-171-718-119 | Sequence 119, App |
| 44 | 6 | 100.0 | 14 | 2 | US-08-173-489C-324 | Sequence 324, App |
| 45 | 6 | 100.0 | 14 | 2 | US-08-232-087A-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1
US-08-859-954-49
Sequence 49, Application US/0885954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramlin
APPLICANT: Hardin, Paul E.
TITLE OF INVENTION: Design and Optimized Primer Library for
TITLE OF INVENTION: Gene Sequencing and Method Thereof
NUMBER OF SEQUENCES: 566
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,954
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligonucleotide"
HYPOTHETICAL: YES
ANTI-SENSE: YES
US-08-859-954-49

Query Match 100.0%; Score 6; DB 3; Length 8;

Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCCTCU 6
 |||||:
Db 3 GCCTCT 8

RESULT 2
US-08-388-353-459/c
Sequence 459, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCCTCU 6
 |||||:
Db 10 GCCTCT 5

RESULT 3
US-08-388-353-460/c
Sequence 460, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCCTCU 6
 |||||:
Db 9 GCCTCT 4

RESULT 4
US-08-388-353-461/c
Sequence 461, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
DB 8 GCCTCT 3

RESULT 5
US-08-388-353-462/c
Sequence 462, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
DB 7 GCCTCT 2

RESULT 6
US-08-388-353-463/c
Sequence 463, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
DB 6 GCCTCT 1

RESULT 7
US-08-488-551B-459/c
Sequence 459, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
DB 10 GCCTCT 5

RESULT 8
US-08-488-551B-460/C
Sequence 460, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
DB 9 GCCTCT 4

RESULT 9
US-08-488-551B-461/C
Sequence 461, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 8 GCCTCT 3

RESULT 10
US-08-488-551B-462/c
Sequence 462, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 7 GCCTCT 2

RESULT 11
US-08-488-551B-463/c
Sequence 463, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 6 GCCTCT 1

RESULT 12
US-09-233-614-2/c
Sequence 2, Application US/092335614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.

APPLICANT: CLARK, CHRISTOPHER L.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-2

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 6 GCCTCT 1

RESULT 13
US-09-235-614-3/C
Sequence 3, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 3
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-3

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 7 GCCTCT 2

RESULT 14
US-09-235-614-4/C
Sequence 4, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 15
US-09-235-614-5/C
Sequence 5, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 5
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-5

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 9 GCCTCT 4

Search completed: May 21, 2003, 08:40:18
Job time : 35.75 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:44:53 ; Search time 65 Seconds
(without alignments)
121.889 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6
Sequence: 1 gccctcu 6

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 828747 seqs, 660231138 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| c 1 | 6 | 100.0 | 9 | US-09-990-186-588 | Sequence 588, App |
| c 2 | 6 | 100.0 | 9 | US-09-990-186-2084 | Sequence 2084, App |
| c 3 | 6 | 100.0 | 9 | US-09-990-186-2086 | Sequence 2086, App |
| c 4 | 6 | 100.0 | 9 | US-09-990-186-2219 | Sequence 2219, App |
| c 5 | 6 | 100.0 | 9 | US-09-990-186-2280 | Sequence 2280, App |
| c 6 | 6 | 100.0 | 9 | US-09-990-186-2471 | Sequence 2471, App |
| c 7 | 6 | 100.0 | 9 | US-09-990-186-2473 | Sequence 2473, App |
| c 8 | 6 | 100.0 | 9 | US-09-989-789-588 | Sequence 588, App |
| c 9 | 6 | 100.0 | 9 | US-09-989-789-2084 | Sequence 2084, App |
| c 10 | 6 | 100.0 | 9 | US-09-989-789-2086 | Sequence 2086, App |
| c 11 | 6 | 100.0 | 9 | US-09-989-789-2219 | Sequence 2219, App |
| c 12 | 6 | 100.0 | 9 | US-09-989-789-2280 | Sequence 2280, App |
| c 13 | 6 | 100.0 | 9 | US-09-989-789-2471 | Sequence 2471, App |
| c 14 | 6 | 100.0 | 9 | US-09-989-789-2473 | Sequence 2473, App |
| c 15 | 6 | 100.0 | 10 | US-09-990-186-617 | Sequence 617, App |
| c 16 | 6 | 100.0 | 10 | US-09-990-186-1279 | Sequence 1279, App |
| c 17 | 6 | 100.0 | 10 | US-09-990-186-1308 | Sequence 1308, App |
| c 18 | 6 | 100.0 | 10 | US-09-990-186-1313 | Sequence 1313, App |
| c 19 | 6 | 100.0 | 10 | US-09-989-789-617 | Sequence 617, App |

| | | | | | |
|------|---|-------|----|---------------------|--------------------|
| c 20 | 6 | 100.0 | 10 | US-09-989-789-1279 | Sequence 1279, App |
| c 21 | 6 | 100.0 | 10 | US-09-989-789-1308 | Sequence 1308, App |
| c 22 | 6 | 100.0 | 10 | US-09-989-789-1313 | Sequence 1313, App |
| c 23 | 6 | 100.0 | 10 | US-10-033-143-153 | Sequence 153, App |
| c 24 | 6 | 100.0 | 10 | US-10-033-143-174 | Sequence 174, App |
| c 25 | 6 | 100.0 | 10 | US-10-033-143-355 | Sequence 355, App |
| c 26 | 6 | 100.0 | 10 | US-10-033-143-566 | Sequence 566, App |
| c 27 | 6 | 100.0 | 10 | US-10-033-143-968 | Sequence 968, App |
| c 28 | 6 | 100.0 | 10 | US-10-033-143-1007 | Sequence 1007, App |
| c 29 | 6 | 100.0 | 10 | US-10-033-143-1120 | Sequence 1120, App |
| c 30 | 6 | 100.0 | 10 | US-10-033-143-1286 | Sequence 1286, App |
| c 31 | 6 | 100.0 | 12 | US-10-033-143-1609 | Sequence 1609, App |
| c 32 | 6 | 100.0 | 11 | US-09-249-155-25 | Sequence 25, App1 |
| c 33 | 6 | 100.0 | 11 | US-09-249-155-207 | Sequence 207, App |
| c 34 | 6 | 100.0 | 11 | US-09-249-155-239 | Sequence 239, App |
| c 35 | 6 | 100.0 | 12 | US-10-131-591A-73 | Sequence 73, App1 |
| c 36 | 6 | 100.0 | 12 | US-09-761-116-1 | Sequence 1, App1 |
| c 37 | 6 | 100.0 | 12 | US-09-384-472-21 | Sequence 21, App1 |
| c 38 | 6 | 100.0 | 12 | US-09-384-472-18 | Sequence 18, App1 |
| c 39 | 6 | 100.0 | 12 | US-09-384-472-21 | Sequence 21, App1 |
| c 40 | 6 | 100.0 | 14 | US-08-591-486B-102 | Sequence 102, App |
| c 41 | 6 | 100.0 | 14 | US-09-578-600-198 | Sequence 198, App |
| c 42 | 6 | 100.0 | 14 | US-09-998-027-26 | Sequence 26, App1 |
| c 43 | 6 | 100.0 | 14 | US-09-504-231A-1431 | Sequence 1431, App |
| c 44 | 6 | 100.0 | 14 | US-09-274-553D-1431 | Sequence 1431, App |
| c 45 | 6 | 100.0 | 15 | US-10-056-414-119 | Sequence 119, App |

ALIGNMENTS

```
RESULT 1
US-09-990-186-588/c
; Sequence 588, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIT, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-588

Query Match          100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCTCU 6
      |||||
Db       8 GCCCTCT 3

RESULT 2
US-09-990-186-2084/c
; Sequence 2084, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIT, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT FILING DATE: 2001-11-20
```

NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2084
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2084

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 3
US-09-990-186-2086/c
Sequence 2086, Application US/09990186
Publication No. US2003006675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2086
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2086

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 4
US-09-990-186-2219/c
Sequence 2219, Application US/09990186
Publication No. US2003006675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2219
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2219

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 5
US-09-990-186-2280/c
Sequence 2280, Application US/09990186
Publication No. US2003006675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2280
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2280

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 6
US-09-990-186-2471/c
Sequence 2471, Application US/09990186
Publication No. US2003006675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2471
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2471

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 7
US-09-990-186-2473/c

```
; Sequence 2473, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-2473
```

```
Query Match          100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 8
US-09-989-789-588/c
; Sequence 588, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-588
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 9
US-09-989-789-2084/c
; Sequence 2084, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 2084
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2084
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 10
US-09-989-789-2086/c
; Sequence 2086, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2086
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2086
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 11
US-09-989-789-2219/c
; Sequence 2219, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2219
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2219
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 12
US-09-989-789-2280/c
; Sequence 2280, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:

; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2280
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA

US-09-989-789-2280

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 13
US-09-989-789-2471/c
; Sequence 2471, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:

; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2471
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA

US-09-989-789-2471

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 14
US-09-989-789-2473/c
; Sequence 2473, Application US/09989789
; Patent No. US20020063379A1

; GENERAL INFORMATION:
; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA

US-09-989-789-2473

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 15
US-09-990-186-617/c
; Sequence 617, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:

; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 617
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA

US-09-990-186-617

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

Search completed: May 21, 2003, 08:44:51
Job time : 65 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:26:13 : Search time 1092.5 Seconds
(without alignments)
88.945 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6
Sequence: 1 gcctcu 6

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_iny:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 100.0 | 15 | 14 | B0511821 | B0511821 EST619236 |
| 2 | 100.0 | 16 | 14 | B0789979 | B0789979 hage0058A |
| 3 | 100.0 | 19 | 9 | A1569191 | A1569191 tr3f112.x |
| 4 | 100.0 | 19 | 17 | A2429998 | A2429998 IM0214F16 |
| 5 | 100.0 | 19 | 17 | A2475079 | A2475079 IM0293B17 |
| 6 | 100.0 | 19 | 17 | A2480905 | A2480905 IM0302N22 |

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 7 | 100.0 | 19 | 17 | A2509071 | A2509071 IM0351A21 |
| 8 | 100.0 | 19 | 17 | A2651803 | A2651803 IM0522N11 |
| 9 | 100.0 | 19 | 17 | A2659603 | A2659603 IM0537N06 |
| 10 | 100.0 | 19 | 17 | A2800646 | A2800646 IM0058H14 |
| 11 | 100.0 | 19 | 17 | A2834038 | A2834038 IM0116H01 |
| 12 | 100.0 | 20 | 14 | B0789776 | B0789776 hage002aA |
| 13 | 100.0 | 20 | 17 | A2309156 | A2309156 IM0013B09 |
| 14 | 100.0 | 20 | 17 | A2366535 | A2366535 IM0115M15 |
| 15 | 100.0 | 20 | 17 | A2615164 | A2615164 IM0444B24 |
| 16 | 100.0 | 20 | 17 | A2797468 | A2797468 IM0053P09 |
| 17 | 100.0 | 20 | 17 | A2807038 | A2807038 IM0069C06 |
| 18 | 100.0 | 21 | 17 | A2303734 | A2303734 IM0003O21 |
| 19 | 100.0 | 21 | 17 | A2317085 | A2317085 IM00035D07 |
| 20 | 100.0 | 21 | 17 | A2342282 | A2342282 IM0075H14 |
| 21 | 100.0 | 21 | 17 | A2650869 | A2650869 IM0521G14 |
| 22 | 100.0 | 22 | 9 | A1023917 | A1023917 ow69F05.s |
| 23 | 100.0 | 22 | 9 | A1128425 | A1128425 qc67f09.x |
| 24 | 100.0 | 22 | 9 | A1660937 | A1660937 wf20b06.x |
| 25 | 100.0 | 22 | 17 | A2787102 | A2787102 IM0033B03 |
| 26 | 100.0 | 22 | 17 | A2792724 | A2792724 IM0045F20 |
| 27 | 100.0 | 22 | 17 | A2803482 | A2803482 IM0063I23 |
| 28 | 100.0 | 22 | 17 | TA204A05P | TA204A05P |
| 29 | 100.0 | 22 | 17 | TA294D03P | TA294D03P |
| 30 | 100.0 | 23 | 17 | A2779967 | A2779967 IM0156C14 |
| 31 | 100.0 | 23 | 17 | A2779967 | A2779967 IM0156C14 |
| 32 | 100.0 | 23 | 17 | A2781980 | A2781980 IM0021N13 |
| 33 | 100.0 | 23 | 17 | A2783504 | A2783504 IM0025L17 |
| 34 | 100.0 | 23 | 17 | A2830526 | A2830526 IM0109M16 |
| 35 | 100.0 | 23 | 17 | A2954682 | A2954682 IM0220P23 |
| 36 | 100.0 | 23 | 17 | A2967993 | A2967993 IM0240J07 |
| 37 | 100.0 | 23 | 17 | BH812514 | BH812514 SALK_0618 |
| 38 | 100.0 | 24 | 17 | A2443047 | A2443047 IM0237I06 |
| 39 | 100.0 | 24 | 17 | A2463313 | A2463313 IM0272L03 |
| 40 | 100.0 | 24 | 17 | A2505865 | A2505865 IM0346C18 |
| 41 | 100.0 | 24 | 17 | A2780307 | A2780307 IM0017J04 |
| 42 | 100.0 | 24 | 17 | A2787390 | A2787390 IM0003J23 |
| 43 | 100.0 | 24 | 17 | A2789936 | A2789936 IM0038L17 |
| 44 | 100.0 | 24 | 17 | A2812679 | A2812679 IM0079H03 |
| 45 | 100.0 | 24 | 17 | A2936903 | A2936903 IM0193E20 |

ALIGNMENTS

RESULT 1
B0511821
LOCUS
DEFINITION
EST619236 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMH018 5' end, mRNA sequence.

ACCESSION
B0511821
VERSION
B0511821.1
KEYWORDS
GI:21370690
SOURCE
EST.
ORGANISM
Solanum tuberosum

REFERENCE
AUTHORS
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., Van der Hoeven, R., Tsai, J. and Karayancheva, S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
Other ESTs: EST619237
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Email: potatoc@tigr.org
This clone is available through the Research Genetics, contact the Research Genetics for further information 1-800-711-6195 or cna@resgen.com
Seq primer: T3.

TITLE
JOURNAL
COMMENT

```

FEATURES
  source
    Location/Qualifiers
      1. .15
        /organism="Solanum tuberosum"
        /cultivar="Kennebec or Bintje"
        /db_xref="taxon:4113"
        /clone="STMH018"
        /clone_1lb="Generation of a set of potato cDNA clones for
        microarray analyses mixed potato tissues"
        /tissue_type="mixed tissues"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
        XhoI; supplier: Combination of untreated and Phytophthora
        infestans-treated libraries of stolons, leaves, leaflets,
        axillary buds of stem explants, petioles, germinating eyes
        , tubers, or roots."

BASE COUNT      3 a      5 c      2 g      5 t
ORIGIN
Query Match      100.0%; Score 6; DB 14; Length 15;
Best Local Similarity 83.3%; Pred. No. 7.8e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCU 6
Db 9 GCCTCT 14

RESULT 2
LOCUS B0789979
DEFINITION B0789979 16 bp mRNA linear EST 30-JUL-2002
             hage005a12 Heterobasidion annosum - Scots pine infection stage
             (HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
             annosum cDNA clone hage005a12, mRNA sequence.
ACCESSION B0789979
VERSION B0789979.1 GI:22004941
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 16)
          Asiegbu,F.O., Nahalkova,J. and Dean,R.A.
          Selected Expressed sequence tags of cDNA clones from the
          interaction of the root rot fungus (Heterobasidion annosum) with
          seedling roots of Scots pine (Pinus sylvestris)
          Unpublished (2001)
JOURNAL Contact: Fred O. Asiegbu
          Dept. of Forest Mycology & Pathology
          Swedish University of Agriculture, Box 7026,S-750 07, Uppsala,
          Sweden
          Tel: +46 18 67 15 98
          Fax: +46 18 30 92 45
          Email: Fred.Asiegbu@mykopat.slu.se
          Seq primer: T7 primer.
FEATURES
  source
    Location/Qualifiers
      1. .16
        /organism="Pinus sylvestris/Heterobasidion annosum"
        /db_xref="taxon:169015"
        /clone="hage005a12"
        /clone_1lb="Heterobasidion annosum - Scots pine infection
        stage (HAGE) subtraction cDNA library"
        /dev_stage="Seedling roots of scots pine were infected for
        6 days with H. annosum"
        /note="Vector: pT-ADV; Site_1: EcoRI; The subtractive
        hybridization cDNA library was constructed from scots pine
        roots infected for 6-days with mycelia of Heterobasidion
        annosum (FP5)."

BASE COUNT      5 a      3 c      3 g      5 t
ORIGIN
Query Match      100.0%; Score 6; DB 14; Length 16;
Best Local Similarity 83.3%; Pred. No. 8e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCU 6
Db 8 GCCTCT 13

RESULT 4
LOCUS A2429998/c
DEFINITION A2429998 19 bp DNA linear GSS 03-OCT-2000
             1M0214F16F Mouse 10Kb plasmid UGCG1M library Mus musculus genomic
             clone UGCG1M0214F16 F, DNA sequence.
ACCESSION A2429998
VERSION A2429998.1 GI:10554011
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 19)

Oy 1 GCCTCU 6
Db 8 GCCTCT 13

RESULT 3
LOCUS A1569191
DEFINITION A1569191 19 bp mRNA linear EST 14-MAY-1999
             tr33fi2.x1 NCI CGAP Panel Homo sapiens cDNA clone IMAGE:2224943 3'
             similar to SW:PCEC_HUMAN P28074 PROTEASOME EPSILON CHAIN PRECURSOR
             , mRNA sequence.
ACCESSION A1569191
VERSION A1569191.1 GI:4532565
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 19)
          NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Life Technologies catalog #: 11548-013
          DNA sequencing by: Washington University Genome Sequencing Center
          clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LMW at:
          www-bio.11nl.gov/db/ftp/image/image.html
          Trace considered overall poor quality
          Insert length: 1117 Std Error: 0.00
          Seq primer: -40bp from G1bco
          High quality sequence stop: 1
          POLYA-No.
FEATURES
  source
    Location/Qualifiers
      1. .19
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:2224943"
        /clone_1lb="NCI-CGAP_Pan1"
        /tissue_type="adenocarcinoma"
        /lab_host="DH10B"
        /note="Organ: pancreas; Vector: pCMV-Sport6; Site_1: SalI;
        Site_2: NotI; Cloned unidirectionally. Primer: 0119g dt.
        Average insert size 1.72 kb. Life Technologies catalog #:
        11548-013"

BASE COUNT      2 a      7 c      5 g      5 t
ORIGIN
Query Match      100.0%; Score 6; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0214 row: F column: 16
Seq primer: CGTTGTAAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UDGCM0214F16"
/clone_1lb="Mouse 10kb plasmid UDGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1/4732114[9b]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
5 a 3 c 9 g 2 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 11 GCCTCT 6

RESULT 5
AZ475079/c 19 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION IM0293B1/F Mouse 10kb plasmid UDGCM library Mus musculus genomic
clone UDGCM0293B1 F, DNA sequence.
ACCESSION AZ475079
VERSION AZ475079.1 GI:10633204
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0293 row: B column: 17
Seq primer: CGTTGTAAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UDGCM0293B17"
/clone_1lb="Mouse 10kb plasmid UDGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1/4732114[9b]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 4 c 9 g 0 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 18 GCCTCT 13

RESULT 6
AZ480905 19 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION IM0302N22R Mouse 10kb plasmid UDGCM library Mus musculus genomic
clone UDGCM0302N22 R, DNA sequence.
ACCESSION AZ480905
VERSION AZ480905.1 GI:10642066
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 22
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0302N22"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
1 a 5 c 9 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||

Db 7 GCCTCT 12

RESULT 7
A2509071/c 19 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0351A21R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCG1M0351A21 R, DNA sequence.

ACCESSION A2509071
VERSION A2509071.1 GI:10690387
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0351 row: A column: 21
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0351A21"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
4 a 4 c 9 g 2 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||

Db 13 GCCTCT 8

RESULT 8
A2651803 19 bp DNA linear GSS 14-DEC-2000
LOCUS A2651803
DEFINITION 1M0522N11R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCG1M0522N11 R, DNA sequence.

ACCESSION A2651803
VERSION A2651803.1 GI:11787672
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0522 row: N column: 11
Seq primer: CACACGAGAACAGCATGACG
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0522N11"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1147321149b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 7 c 5 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

DB 7 GCCTCT 12

RESULT 9
AZ659603/c 19 bp DNA linear GSS 14-DEC-2000
LOCUS
DEFINITION 1M037N06F Mouse 10kb plasmid UGGCM library Mus musculus genomic
clone UGGCM0537N06 F, DNA sequence.
ACCESSION AZ659603
VERSION AZ659603.1 GI:11796749
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0537 row: N column: 06
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0537N06"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1147321149b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
7 a 2 c 7 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

DB 12 GCCTCT 7

RESULT 10
AZ800646/c 19 bp DNA linear GSS 16-FEB-2001
LOCUS
DEFINITION 2M0058H14R Mouse 10kb plasmid UGGCM library Mus musculus genomic
clone UGGCM0058H14 R, DNA sequence.
ACCESSION AZ800646
VERSION AZ800646.1 GI:12952969
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts.

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0058 row: H column: 14
Seq primer: CACACAGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0058H14"
/clone_1lb="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
3 a 4 c 7 g 5 t

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 7 GCCTCT 2

RESULT 11
AZ834038/c 19 bp DNA linear GSS 20-FEB-2001
LOCUS 2M0116H01R Mouse 10kb plasmid U06C1M library Mus musculus genomic
DEFINITION clone U06C2M0116H01 R, DNA sequence.
ACCESSION AZ834038
VERSION AZ834038.1 GI:13003946
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts.

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0116 row: H column: 01
Seq primer: CACACAGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0116H01"
/clone_1lb="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
6 a 2 c 7 g 4 t

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 6 GCCTCT 1

RESULT 12
BQ789776 20 bp mRNA linear EST 30-JUL-2002
LOCUS BQ789776
DEFINITION hage002a10 Heterobasidion annosum - Scots pine infection stage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage002a10, mRNA sequence.
ACCESSION BQ789776
VERSION BQ789776.1 GI:22004738
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 20)

AUTHORS Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
TITLE Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (*Heterobasidion annosum*) with seedling roots of Scots pine (*Pinus sylvestris*)
JOURNAL Unpublished (2001)
COMMENT Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden

Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer.

FEATURES
source

1.20
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone="hage002a10"
/clone_1lb="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
/note="Vector: pT-Adv; Site: 1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of *Heterobasidion annosum* (FP5)."
BASE COUNT 4 a 5 c 3 g 7 t 1 others
ORIGIN

Query Match 100.0%; Score 6; DB 14; Length 20;
Best Local Similarity 83.3%; Pred. No. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 12 GCCTCT 17

RESULT 13
AZ309156
LOCUS 20 bp DNA linear GSS 29-SEP-2000
DEFINITION IM0013B09F Mouse 10kb plasmid UGCM library Mus musculus genomic
ACCESSION clone UGCM0013B09 F, DNA sequence.
VERSION AZ309156.1 GI:10349862
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0013 row: B column: 09
Seq primer: CGTTGTAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 20.
location/Qualifiers
1..20
source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCM0013B09"
/clone_1lb="Mouse 10kb plasmid UGCM library"
/sex="Male"
/lab_host="E. coli strain XL10-gold, T1-resistant, F-"
/note="Vector: pMD29v; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114[gb|AF129072.1]) a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 11 c 4 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 3 GCCTCT 8

RESULT 14
AZ366535
LOCUS 20 bp DNA linear GSS 02-OCT-2000
DEFINITION IM0115M15R Mouse 10kb plasmid UGCM library Mus musculus genomic
ACCESSION clone UGCM00115M15 R, DNA sequence.
VERSION AZ366535.1 GI:10480235
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0115 row: M column: 15
Seq primer: CACACAGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
location/Qualifiers
1..20
source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UNGC1M0115M15"
/clone_1lb="Mouse 10kb plasmid UNGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      5 a      6 c      3 g      6 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. NO. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GCCTCU 6
        |||||:
        14 GCCTCT 19

```

```

RESULT 15
LOCUS      A2615164
DEFINITION 20 bp DNA linear GSS 13-DEC-2000
            A2615164
            clone UNGC1M044B24 F, DNA sequence.
ACCESSION  A2615164
VERSION     A2615164
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE  1 (bases 1 to 20)
AUTHORS    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Irlam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: cdunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0444 row: B column: 24
            Seq primer: CGTTGTAAACGACGCCACT
            Class: plasmid ends
            High quality sequence stop: 20.

```

```

FEATURES
Source      I. 20
Location/Qualifiers

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UNGC1M044B24"
/clone_1lb="Mouse 10kb plasmid UNGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      3 a      7 c      5 g      5 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. NO. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GCCTCU 6
        |||||:
        DB      7 GCCTCT 12

```

```

Search completed: May 21, 2003, 08:37:45
Job time : 1092.5 secs

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:56:27 ; Search time 333.5 Seconds

(without alignments)
323.369 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| C 1 | 6 | 100.0 | 10 | 6 | AR128987 | AR128987 Sequence |
| C 2 | 6 | 100.0 | 10 | 6 | AR128988 | AR128988 Sequence |
| C 3 | 6 | 100.0 | 10 | 6 | AR128989 | AR128989 Sequence |
| C 4 | 6 | 100.0 | 10 | 6 | AR128990 | AR128990 Sequence |
| C 5 | 6 | 100.0 | 10 | 6 | AR128991 | AR128991 Sequence |
| C 6 | 6 | 100.0 | 10 | 6 | AX152349 | AX152349 Sequence |
| C 7 | 6 | 100.0 | 10 | 6 | AX152410 | AX152410 Sequence |
| C 8 | 6 | 100.0 | 10 | 6 | AX152849 | AX152849 Sequence |
| C 9 | 6 | 100.0 | 10 | 6 | AX152886 | AX152886 Sequence |
| C 10 | 6 | 100.0 | 10 | 6 | AX153384 | AX153384 Sequence |
| C 11 | 6 | 100.0 | 10 | 6 | AX153513 | AX153513 Sequence |
| C 12 | 6 | 100.0 | 10 | 6 | E16890 | E16890 DNA sequence |
| C 13 | 6 | 100.0 | 11 | 6 | A91501 | A91501 Sequence 28 |
| C 14 | 6 | 100.0 | 11 | 6 | A91503 | A91503 Sequence 30 |
| C 15 | 6 | 100.0 | 11 | 6 | AX470495 | AX470495 Sequence |
| C 16 | 6 | 100.0 | 11 | 6 | AX470740 | AX470740 Sequence |
| C 17 | 6 | 100.0 | 11 | 6 | AX470747 | AX470747 Sequence |
| C 18 | 6 | 100.0 | 11 | 6 | AX471502 | AX471502 Sequence |
| C 19 | 6 | 100.0 | 11 | 6 | AX471630 | AX471630 Sequence |
| C 20 | 6 | 100.0 | 11 | 6 | AX471677 | AX471677 Sequence |
| C 21 | 6 | 100.0 | 11 | 6 | AX471805 | AX471805 Sequence |
| C 22 | 6 | 100.0 | 12 | 6 | A06058 | A06058 Synthetic P |
| C 23 | 6 | 100.0 | 12 | 6 | A06059 | A06059 Synthetic P |
| C 24 | 6 | 100.0 | 12 | 6 | A15123 | A15123 Nucleotide |
| C 25 | 6 | 100.0 | 12 | 6 | A16601 | A16601 Nucleotide |
| C 26 | 6 | 100.0 | 12 | 6 | A16602 | A16602 Nucleotide |
| C 27 | 6 | 100.0 | 12 | 6 | A47656 | A47656 Sequence 16 |
| C 28 | 6 | 100.0 | 12 | 6 | A61520 | A61520 Sequence 89 |
| C 29 | 6 | 100.0 | 12 | 6 | A91489 | A91489 Sequence 16 |
| C 30 | 6 | 100.0 | 12 | 6 | AR027874 | AR027874 Sequence |
| C 31 | 6 | 100.0 | 12 | 6 | AR075465 | AR075465 Sequence |
| C 32 | 6 | 100.0 | 12 | 6 | AR075468 | AR075468 Sequence |
| C 33 | 6 | 100.0 | 12 | 6 | AR101000 | AR101000 Sequence |
| C 34 | 6 | 100.0 | 12 | 6 | AR137925 | AR137925 Sequence |
| C 35 | 6 | 100.0 | 12 | 6 | AR153916 | AR153916 Sequence |
| C 36 | 6 | 100.0 | 12 | 6 | AR153919 | AR153919 Sequence |
| C 37 | 6 | 100.0 | 12 | 6 | AR167701 | AR167701 Sequence |
| C 38 | 6 | 100.0 | 12 | 6 | AR178533 | AR178533 Sequence |
| C 39 | 6 | 100.0 | 12 | 6 | AR178536 | AR178536 Sequence |
| C 40 | 6 | 100.0 | 12 | 6 | AR199100 | AR199100 Sequence |
| C 41 | 6 | 100.0 | 12 | 6 | AX233626 | AX233626 Sequence |
| C 42 | 6 | 100.0 | 12 | 6 | E29585 | E29585 Method for |
| C 43 | 6 | 100.0 | 12 | 6 | E38691 | E38691 Method for |
| C 44 | 6 | 100.0 | 12 | 6 | E64117 | E64117 Method for |
| C 45 | 6 | 100.0 | 12 | 6 | I07918 | I07918 Sequence 30 |

ALIGNMENTS

RESULT 1
AR128987/c
LOCUS AR128987 10 bp DNA linear PART 16-MAY-2001
DEFINITION Sequence 2 from patent US 6183966.
ACCESSION AR128987
VERSION AR128987.1 GI:14116649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 2 06-FEB-2001.

FEATURES Location/Qualifiers
source 1.10
/organism="unknown"
BASE COUNT 2 a 3 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
Db 6 GCCCTC 1

RESULT 2
LOCUS ARI28988 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6183966.
ACCESSION ARI28988
VERSION ARI28988.1 GI:14116650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 3 06-FEB-2001;
FEATURES Location/Qualifiers
source 1.10
/organism="unknown"

BASE COUNT 2 a 3 c 5 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
Db 7 GCCCTC 2

RESULT 3
LOCUS ARI28989 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 4 from patent US 6183966.
ACCESSION ARI28989
VERSION ARI28989.1 GI:14116651
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 4 06-FEB-2001;
FEATURES Location/Qualifiers
source 1.10
/organism="unknown"

BASE COUNT 3 a 3 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
Db 8 GCCCTC 3

RESULT 4
LOCUS ARI28990 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6183966.
ACCESSION ARI28990
VERSION ARI28990.1 GI:14116652
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 5 06-FEB-2001;
FEATURES Location/Qualifiers
source 1.10
/organism="unknown"

BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
Db 9 GCCCTC 4

RESULT 5
LOCUS ARI28991 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6183966.
ACCESSION ARI28991
VERSION ARI28991.1 GI:14116653
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 6 06-FEB-2001;
FEATURES Location/Qualifiers
source 1.10
/organism="unknown"

BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
Db 10 GCCCTC 5

RESULT 6
LOCUS AX152349 10 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 264 from Patent W00138577.
ACCESSION AX152349
VERSION AX152349.1 GI:14534000
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 264 31-MAY-2001;
          The Johns Hopkins University (US)
FEATURES
    source
        1. .10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      0 a      5 c      2 g      3 t
ORIGIN

Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTCU 6
        |||||:
        1 GCCTCT 6

RESULT 7
AX152410
LOCUS AX152410
DEFINITION Sequence 325 from Patent W00138577.
ACCESSION AX152410
VERSION AX152410.1 GI:14534061
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 325 31-MAY-2001;
          The Johns Hopkins University (US)
FEATURES
    source
        1. .10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      1 a      5 c      2 g      2 t
ORIGIN

Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTCU 6
        |||||:
        1 GCCTCT 6

RESULT 8
AX152849
LOCUS AX152849
DEFINITION Sequence 764 from Patent W00138577.
ACCESSION AX152849
VERSION AX152849.1 GI:14534500
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 764 31-MAY-2001;
          The Johns Hopkins University (US)
FEATURES
    source
        1. .10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      0 a      5 c      2 g      3 t
ORIGIN

Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTCU 6
        |||||:
        1 GCCTCT 6

RESULT 9
AX152886/c
LOCUS AX152886/c
DEFINITION Sequence 801 from Patent W00138577.
ACCESSION AX152886
VERSION AX152886.1 GI:14534537
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 801 31-MAY-2001;
          The Johns Hopkins University (US)
FEATURES
    source
        1. .10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      5 a      1 c      3 g      1 t
ORIGIN

Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTCU 6
        |||||:
        8 GCCTCT 3

RESULT 10
AX153384
LOCUS AX153384
DEFINITION Sequence 1299 from Patent W00138577.
ACCESSION AX153384
VERSION AX153384.1 GI:14535035
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 1299 31-MAY-2001;
          The Johns Hopkins University (US)
FEATURES
    source
        1. .10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      0 a      4 c      3 g      3 t
ORIGIN

Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTCU 6
        |||||:
        1 GCCTCT 6
```

Db 2 GCCTCT 7

RESULT 11
LOCUS AX153513 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1428 from Patent WO0138577.
ACCESSION AX153513
VERSION AX153513.1 GI:14535164
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 10)
TITLE Velculescu, V.E., Vogelstein, B. and Kinzler, K.W.
JOURNAL Human transcriptomes
Patent: WO 0138577-A 1428 31-MAY-2001;
The Johns Hopkins University (US)
Location/Qualifiers
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 0 a 4 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCT 6
|||||
Db 1 GCCTCT 6

RESULT 12
E16890 10 bp DNA linear PAT 28-JUL-1999
LOCUS E16890
DEFINITION DNA sequence required for efficient protein transcription in
Brevibacterium flavum.
ACCESSION E16890
VERSION E16890.1 GI:5711573
KEYWORDS JP 1998229881-A/31.
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS 1 (bases 1 to 10)
TITLE Kobayashi, M., Man, T. and Yugawa, H.
DNA HAVING SEQUENCE CAPABLE OF EFFICIENTLY TRANSLATING PROTEIN IN
CORYNEFORM BACTERIA
Patent: JP 1998229881-A 31 02-SEP-1998;
MITSUBISHI CHEM CORP
JOURNAL
COMMENT
OS Brevibacterium flavum
PN JP 1998229881-A/31
PD 02-SEP-1998
PF 19-FEB-1997 JP 1997035338
PI KOBAYASHI MIKI, MAN TOMOKO, YUGAWA HIDEAKI
PC C12N15/09,C07H21/04,C12N1/21//C12N9/38,C12Q1/68,(C12N15/09, PC
C12N1/19),
PC (C12N1/21,C12N1/13),(C12N9/38,C12N1/19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH key
FT source 1..10
/organism="Brevibacterium flavum" FT
/strain="MJ-233",
Location/Qualifiers
FEATURES
source 1..10
/organism="Corynebacterium glutamicum"

BASE COUNT 3 a 1 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCT 6
|||||
Db 9 GCCTCT 4

RESULT 13
A91501 11 bp DNA linear PAT 22-JAN-2000
LOCUS A91501
DEFINITION Sequence 28 from Patent WO9824928.
ACCESSION A91501
VERSION A91501.1 GI:6740456
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
AUTHORS 1 (bases 1 to 11)
TITLE Pallsgaard, N. and Hokland, P.
JOURNAL DETECTION OF CHROMOSOMAL ABNORMALITIES
Patent: WO 9824928-A 28 11-JUN-1998;
PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
FEATURES
source 1..11
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1 a 5 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2e+06; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCT 6
|||||
Db 5 GCCTCT 10

RESULT 14
A91503 11 bp DNA linear PAT 22-JAN-2000
LOCUS A91503
DEFINITION Sequence 30 from Patent WO9824928.
ACCESSION A91503
VERSION A91503.1 GI:6740458
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
AUTHORS 1 (bases 1 to 11)
TITLE Pallsgaard, N. and Hokland, P.
JOURNAL DETECTION OF CHROMOSOMAL ABNORMALITIES
Patent: WO 9824928-A 30 11-JUN-1998;
PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
FEATURES
source 1..11
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 0 a 5 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2e+06; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCT 6
|||||
Db 5 GCCTCT 10

RESULT 15
AX470495/c 11 bp DNA linear PAT 09-AUG-2002
LOCUS AX470495
DEFINITION Sequence 72 from Patent WO02053773.
ACCESSION AX470495
VERSION AX470495.1 GI:22205620
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hofmann, K., Conradt, M. and Petersohn, D.
TITLE Method for determining skin stress or skin ageing in vitro
JOURNAL Patent. WO 02053773-A 72 11-JUL-2002;
HENKEL KGAA (DE)
FEATURES
source location/Qualifiers
1. 11
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 6 a. 2 c 3 g 0 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3
Search completed: May 21, 2003, 05:12:46
Job time : 336 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:52:32 ; Search time 156 Seconds

(without alignments)
86.615 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 6 | 100.0 | 8 | 20 | AA29578 |
| 2 | 6 | 100.0 | 8 | 21 | AAA80736 |
| 3 | 6 | 100.0 | 9 | 24 | ABQ71469 |
| 4 | 6 | 100.0 | 9 | 24 | ABQ71786 |
| 5 | 6 | 100.0 | 9 | 24 | ABQ71788 |
| 6 | 6 | 100.0 | 9 | 24 | ABQ71921 |
| 7 | 6 | 100.0 | 9 | 24 | ABQ71982 |
| 8 | 6 | 100.0 | 9 | 24 | ABQ72173 |
| 9 | 6 | 100.0 | 9 | 24 | ABQ72175 |

| | | | | | | |
|------|---|-------|----|----|----------|--------------------|
| c 10 | 6 | 100.0 | 10 | 16 | AAQ96863 | HIV-1 NL4-3 nef ge |
| c 11 | 6 | 100.0 | 10 | 16 | AAQ96864 | HIV-1 NL4-3 nef ge |
| c 12 | 6 | 100.0 | 10 | 16 | AAQ96865 | HIV-1 NL4-3 nef ge |
| c 13 | 6 | 100.0 | 10 | 16 | AAQ96866 | HIV-1 NL4-3 nef ge |
| c 14 | 6 | 100.0 | 10 | 16 | AAQ96867 | HIV-1 NL4-3 nef ge |
| c 15 | 6 | 100.0 | 10 | 20 | AAQ86201 | SAGE tag used to 1 |
| c 16 | 6 | 100.0 | 10 | 20 | AAQ00518 | Antisense oligonuc |
| c 17 | 6 | 100.0 | 10 | 20 | AAQ00519 | Antisense oligonuc |
| c 18 | 6 | 100.0 | 10 | 20 | AAQ00522 | Antisense oligonuc |
| c 19 | 6 | 100.0 | 10 | 20 | AAQ00520 | Antisense oligonuc |
| c 20 | 6 | 100.0 | 10 | 20 | AAQ00521 | Antisense oligonuc |
| c 21 | 6 | 100.0 | 10 | 21 | AAQ56532 | Human macrophage 9 |
| c 22 | 6 | 100.0 | 10 | 21 | AAQ77725 | Human dendritic ce |
| c 23 | 6 | 100.0 | 10 | 21 | AAQ77726 | Human dendritic ce |
| c 24 | 6 | 100.0 | 10 | 21 | AAQ77927 | Human dendritic ce |
| c 25 | 6 | 100.0 | 10 | 21 | AAQ78138 | Human dendritic ce |
| c 26 | 6 | 100.0 | 10 | 21 | AAQ78540 | Human dendritic ce |
| c 27 | 6 | 100.0 | 10 | 21 | AAQ78579 | Human dendritic ce |
| c 28 | 6 | 100.0 | 10 | 21 | AAQ78692 | Human dendritic ce |
| c 29 | 6 | 100.0 | 10 | 21 | AAQ78858 | Human dendritic ce |
| c 30 | 6 | 100.0 | 10 | 21 | AAQ79181 | Human dendritic ce |
| c 31 | 6 | 100.0 | 10 | 21 | AAQ79724 | Human colon tumour |
| c 32 | 6 | 100.0 | 10 | 21 | AAQ79825 | Human lung prefe |
| c 33 | 6 | 100.0 | 10 | 21 | AAQ79884 | Human dendritic ce |
| c 34 | 6 | 100.0 | 10 | 21 | AAQ81089 | Metastatic breast |
| c 35 | 6 | 100.0 | 10 | 21 | AAQ81276 | Metastatic breast |
| c 36 | 6 | 100.0 | 10 | 21 | AAQ81485 | Metastatic breast |
| c 37 | 6 | 100.0 | 10 | 21 | AAQ81751 | Metastatic breast |
| c 38 | 6 | 100.0 | 10 | 21 | AAQ81798 | Metastatic breast |
| c 39 | 6 | 100.0 | 10 | 21 | AAQ82317 | Metastatic breast |
| c 40 | 6 | 100.0 | 10 | 21 | AAQ82317 | Metastatic breast |
| c 41 | 6 | 100.0 | 10 | 21 | AAQ82473 | Metastatic breast |
| c 42 | 6 | 100.0 | 10 | 21 | AAQ82783 | Metastatic breast |
| c 43 | 6 | 100.0 | 10 | 21 | AAQ82938 | Metastatic breast |
| c 44 | 6 | 100.0 | 10 | 21 | AAQ83462 | Metastatic breast |
| c 45 | 6 | 100.0 | 10 | 21 | AAQ84090 | Metastatic breast |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| ID | AA29578 |
| AA29578 | standard; DNA: 8 BP. |
| AC | AA29578; |
| XX | |
| DT | 03-JUN-1999 (first entry) |
| XX | |
| DE | Primer for human G-protein coupled receptor genes. |
| XX | |
| KW | Nucleic acid amplification; nuclear receptor; G-protein coupled receptor; |
| KW | apoptosis; DNA repair; DNA replication; plant biology; agriculture; |
| KW | human; veterinary medicine; reproduction; microbiology; hybridisation; |
| KW | environmental science; DNA fingerprinting; PCR primer; ss. |
| XX | |
| OS | Synthetic. |
| OS | Homo sapiens. |
| XX | |
| PN | WO9911823-A2. |
| XX | |
| PD | 11-MAR-1999. |
| XX | |
| PF | 04-SEP-1998; 98WO-US18392. |
| XX | |
| PR | 05-SEP-1997; 97US-0925816. |
| XX | |
| PA | (KIMM-) KIMMEL CANCER CENT SYDNEY. |
| XX | |
| PI | McClelland M, Pesole G; |
| XX | |
| DR | WPI; 1999-205200/17. |
| XX | |

PT Subset of primers able to amplify group of related sequences
 XX
 PS Claim 21; Page 75; 92pp; English.
 XX
 CC The invention provides primers (AAx29501-x29679) for identifying
 CC sequences encoding structurally or functionally related proteins such as
 CC nuclear or G-protein coupled receptors, apoptosis-related or DNA
 CC repair/replication proteins. The identified sequences are broadly useful
 CC in plant biology, agriculture, human or veterinary medicine,
 CC reproduction, microbiology or environmental science, e.g. to study
 CC expression of nuclear receptors at different stages of tissue development
 CC or after treatment with particular drugs. It is also used for DNA
 CC fingerprinting (to generate products useful for differential
 CC hybridisation), or, where a 3'-anchor primer is used, to isolate the
 CC 3'-ends of mRNA sequences. Sequences AAx29526-x29583 represent claimed
 CC primers specific for human G-protein coupled receptor genes.
 XX
 SQ Sequence 8 BP; 0 A; 4 C; 2 G; 2 T; 0 other:
 Query Match 100.0%; Score 6; DB 20; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.7e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCT 6
 Db 1 GCCTCT 6
 RESULT 2
 AAA80736
 ID AAA80736 standard; DNA: 8 BP.
 AC
 XX AAA80736;
 XX
 DT 24-NOV-2000 (first entry)
 DE A. thaliana primer walking octamer SEQ ID NO: 49.
 XX
 KW -Primer walking; octamer; primer; DNA sequencing; PCR; ss.
 OS Arabidopsis thaliana.
 XX
 PN US6083695-A.
 PD 04-JUL-2000.
 XX
 PF 21-MAY-1997; 97US-0859954.
 XX
 PR 15-APR-1996; 96US-0632782.
 XX
 PA (UYHO-) UNIV HOUSTON.
 XX (HARD/) HARDIN S H.
 XX
 PI Hardin PE, Hardin SH, Homayouni R;
 XX
 DR WPI; 2000-474852/41.
 XX
 PT Sequencing an unknown DNA molecule for the polymerase chain reaction
 PT and other primer processes comprises primer walking of octamer
 PT oligonucleotides -
 XX
 PS Example 8; Column 49-50; 161pp; English.
 XX
 CC This invention describes a novel method for sequencing an unknown DNA
 CC molecule which comprises selecting a library primer from an octamer
 CC oligonucleotide library consisting of 48 8-bp sequences and
 CC corresponding complementary sequences, where the library primer is
 CC complementary to a known sequence adjacent to the unknown sequence or
 CC is complementary to a sequence in a known extension product. The method
 CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
 CC which make use of primers. The octamers are used to identify coding
 CC sequences. Primer walking using the octamer libraries is advantageous
 CC over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
 CC directed and methodical approach. AAA80688-A81253 represent the octamer
 CC primers used in the primer walking method of the invention.
 XX
 SQ Sequence 8 BP; 0 A; 3 C; 1 G; 4 T; 0 other:
 Query Match 100.0%; Score 6; DB 21; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.7e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCT 6
 Db 3 GCCTCT 8
 RESULT 3
 AB071469/C
 ID AB071469 standard; DNA: 9 BP.
 AC
 XX AB071469;
 XX
 DT 28-AUG-2002 (first entry)
 DE Zinc finger protein related oligonucleotide target SEQ ID NO:588.
 XX
 DE Zinc finger protein related oligonucleotide target SEQ ID NO:588.
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200242459-A2.
 PD 30-MAY-2002.
 XX
 PF 20-NOV-2001; 2001WO-US43438.
 XX
 PR 20-NOV-2000; 2000US-0716637.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Liu Q;
 XX
 DR WPI; 2002-500284/53.
 XX
 PT New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering.
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 XX
 PS Example 1; Page 44; 81pp; English.
 XX
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (i) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 CC that it binds to the S2 target subsite, and selecting the F3 zinc
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. AB071213 to AB072214 and AB448191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;
SQ Best Local Similarity 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 4
ABQ71786/c
ID ABQ71786 standard; DNA: 9 BP.
AC ABQ71786;
DT 28-AUG-2002 (first entry)
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2084.
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX Homo sapiens.
OS Synthetic.
OS
PN MO200242459-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001MO-US43438.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Liu Q;
XX
DR WPI; 2002-500284/53.
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
PS Example 1; Page 55; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

QY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 5
ABQ71788/c
ID ABQ71788 standard; DNA: 9 BP.
AC ABQ71788;
DT 28-AUG-2002 (first entry)
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2086.
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX Homo sapiens.
OS Synthetic.
OS
PN MO200242459-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001MO-US43438.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Liu Q;
XX
DR WPI; 2002-500284/53.
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
PS Example 1; Page 55; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
 |||||:
 Db 8 GCCTCT 3

RESULT 6

ABQ71921/C
 ID ABQ71921 standard; DNA; 9 BP.

AC ABQ71921;

DT 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2219.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN W0200242459-A2.

XX 30-MAY-2002.

PF 20-NOV-2001; 2001WO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI L1u Q:

DR WPI: 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
 gene function and for human therapeutics and plant engineering,
 comprises first, second and third zinc fingers, ordered from N- to
 C-terminus

XX Example 1; Page 58; 81bp; English.

CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target site. Also described are: (I) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target site, selecting the F2 zinc finger such
 CC that it binds to the S2 target site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sites having the nucleotide G in the 5'-most position of the
 CC substrate. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.4e+08;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
 |||||:

Db 8 GCCTCT 3

RESULT 7

ABQ71982/C
 ID ABQ71982 standard; DNA; 9 BP.

AC ABQ71982;

DT 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2280.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN W0200242459-A2.

XX 30-MAY-2002.

PF 20-NOV-2001; 2001WO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI L1u Q:

DR WPI: 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
 gene function and for human therapeutics and plant engineering, to
 comprises first, second and third zinc fingers, ordered from N- to
 C-terminus

XX Example 1; Page 59; 81bp; English.

CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target site. Also described are: (I) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target site, selecting the F2 zinc finger such
 CC that it binds to the S2 target site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sites having the nucleotide G in the 5'-most position of the
 CC substrate. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.4e+08;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
 |||||:
 Db 8 GCCTCT 3

RESULT 8
AB072173/c
ID AB072173 standard; DNA: 9 BP.
XX
XX AB072173;
XX
XX 28-AUG-2002 (first entry)
XX
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2471.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX Homo sapiens.
XX OS
XX Synthetic.
XX PN MO200242459-A2.
XX PD 30-MAY-2002.
XX
XX 20-NOV-2001; 2001MO-US43438.
XX PF
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX PA
XX
XX
XX Liu Q;
XX
XX WPI; 2002-500284/53.
XX
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
XX
XX Example 1; Page 63; 81bp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsequence. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsequence, selecting the F2 zinc finger such
CC that it binds to the S2 target subsequence, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsequence, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsequences having the nucleotide G in the 5'-most position of the
CC target site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. AB072173 to AB072214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2,4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
| | | | |
| | | | |
Db 8 GCCTCT 3

RESULT 9
AB072175/c
ID AB072175 standard; DNA: 9 BP.

XX
XX AB072175;
AC
XX
XX 28-AUG-2002 (first entry)
XX
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2473.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX Homo sapiens.
XX OS
XX Synthetic.
XX PN MO200242459-A2.
XX PD 30-MAY-2002.
XX
XX 20-NOV-2001; 2001MO-US43438.
XX PF
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX PA
XX
XX
XX Liu Q;
XX
XX WPI; 2002-500284/53.
XX
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
XX
XX Example 1; Page 63; 81bp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsequence. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsequence, selecting the F2 zinc finger such
CC that it binds to the S2 target subsequence, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsequence, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsequences having the nucleotide G in the 5'-most position of the
CC target site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. AB072173 to AB072214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2,4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
| | | | |
| | | | |
Db 8 GCCTCT 3

RESULT 10
AA096863/c
ID AA096863 standard; DNA: 10 BP.
XX
XX AA096863;
XX

```

DT 26-MAR-1996 (first entry)
XX
XX HIV-1 NL4-3 nef gene nucleotide deletion 458.
DE
XX
XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX W09521912-A1.
PN
XX
XX 17-AUG-1995.
PD
XX
XX 14-FEB-1995; 95WO-AU00063.
PF
XX
XX 23-DEC-1994; 94AU-0000284.
PR 14-FEB-1994; 94AU-0003864.
PR 21-FEB-1994; 94AU-0004002.
XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
PS
XX
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AAQ96406-Q97018) from the nef gene and/or
CC 1 or more decanucleotides (AAQ97019-Q97166) from the LTR region; the
CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
CC (AAQ96141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
XX Sequence 10 BP; 4 A; 1 C; 4 G; 1 T; 0 other;
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
Db 10 GCCTCT 5

```

```

XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
PS
XX
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AAQ96406-Q97018) from the nef gene and/or
CC 1 or more decanucleotides (AAQ97019-Q97166) from the LTR region; the
CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene
CC (AAQ96141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
XX Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
Db 9 GCCTCT 4

```

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene (AA096141). The resulting avirulent HIV strains are still capable of inducing an immune response in humans, and enable the generation of CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 13
AA096866/c
ID AA096866 standard; DNA: 10 BP.

XX AA096866;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 461.

KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

PD 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

DR WPI; 1995-293115/38.

PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or LTR region - can be used in a vaccine to inhibit/reduce productive infection in an individual by a pathogenic strain

PS Claim 13; Page 194; 301pp; English.

CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or more deancucleotides (AA096406-097018) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 5 A; 2 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
DB 7 GCCTCT 2

RESULT 14

AA096867/c
ID AA096867 standard; DNA: 10 BP.

XX AA096867;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 462.

KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

PD 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

DR WPI; 1995-293115/38.

PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or LTR region - can be used in a vaccine to inhibit/reduce productive infection in an individual by a pathogenic strain

PS Claim 13; Page 194; 301pp; English.

CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or more deancucleotides (AA096406-097018) from the nef gene and/or

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 4 A; 2 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
DB 6 GCCTCT 1

RESULT 15

AA096201
ID AA096201 standard; DNA: 10 BP.

XX AA096201;

DT 22-SEP-1999 (first entry)

DE SAGE tag used to identify transcripts which are enhanced by p53.

KW p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity; neoplastic; p53 binding site; PIG-3 promoter; SAGE tag; ss.

XX Homo sapiens.

XX WO9914356-A2.

XX 25-MAR-1999.

```

XX 17-SEP-1998; 98WO-US19300.
PF 30-MAR-1998; 98US-0079817.
XX 17-SEP-1997; 97US-0059153.
PR
XX (UT0 ) UNIV JOHNS HOPKINS.
PA
XX Kinzler KW, Polyak K, Vogelstein B;
PI WPI: 1999-443793/37.
XX
DR Use of p53 transcription tags to determine p53 status in, e.g.
XX cancer diagnosis
PT
XX Example 1: Page 25; 73pp; English.
PS
XX The specification describes the use of p53 transcription tags for
CC developing products to determine p53 status, to diagnose cancer
CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
CC A method for diagnosing cancer or determining p53 status in a sample
CC suspected for being neoplastic comprises comparing the level of
CC transcription of an RNA transcript in a first sample (s1) of a first
CC tissue (t1) to the level of transcription of the transcript in a second
CC sample (s2) of a second tissue (s2), where s1 is suspected of being
CC neoplastic and s2 is a normal human tissue (of the same type) and the
CC transcript is identified by a tag; and categorizing s1 as neoplastic
CC or as having a mutant p53 when transcription is found to be the same
CC or lower in the first, than in s2. The methods and products can be used
CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
CC or carcinogenicity of a test agent. AAX86201-33 represent SAGE tags
CC used to identify transcripts which are enhanced by p53.
XX
SQ Sequence 10 BP; 0 A; 6 C; 1 G; 3 T; 0 other;

```

```

Query Match 100.0%; Score 6; DB 20; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

```

```

OY 1 GCCTCU 6
   |||||
DB 4 GCCTCT 9

```

Search completed: May 21, 2003, 04:50:10
 Job time : 156 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:37:42 ; Search time 35.25 Seconds

(without alignments)
52.200 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6

Sequence: 1 gcctca 6

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A_COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B_COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCOTUS_COMB.seq:*
6: /cgn2_6/prodata/2/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------------------|-------------------|
| 1 | 6 | 100.0 | 8 | 3 | US-08-859-954-49 | Sequence 49, App1 |
| 2 | 6 | 100.0 | 10 | 3 | US-08-388-353-459 | Sequence 459, App |
| 3 | 6 | 100.0 | 10 | 3 | US-08-388-353-460 | Sequence 460, App |
| 4 | 6 | 100.0 | 10 | 3 | US-08-388-353-461 | Sequence 461, App |
| 5 | 6 | 100.0 | 10 | 3 | US-08-388-353-462 | Sequence 462, App |
| 6 | 6 | 100.0 | 10 | 3 | US-08-388-353-463 | Sequence 463, App |
| 7 | 6 | 100.0 | 10 | 3 | US-08-488-551B-459 | Sequence 459, App |
| 8 | 6 | 100.0 | 10 | 3 | US-08-488-551B-460 | Sequence 460, App |
| 9 | 6 | 100.0 | 10 | 3 | US-08-488-551B-461 | Sequence 461, App |
| 10 | 6 | 100.0 | 10 | 3 | US-08-488-551B-462 | Sequence 462, App |
| 11 | 6 | 100.0 | 10 | 3 | US-08-488-551B-463 | Sequence 463, App |
| 12 | 6 | 100.0 | 10 | 4 | US-09-235-614-2 | Sequence 2, App1 |
| 13 | 6 | 100.0 | 10 | 4 | US-09-235-614-3 | Sequence 3, App1 |
| 14 | 6 | 100.0 | 10 | 4 | US-09-235-614-4 | Sequence 4, App1 |
| 15 | 6 | 100.0 | 10 | 4 | US-09-235-614-5 | Sequence 5, App1 |
| 16 | 6 | 100.0 | 10 | 4 | US-09-235-614-6 | Sequence 6, App1 |
| 17 | 6 | 100.0 | 10 | 4 | US-09-154-750A-4 | Sequence 4, App1 |
| 18 | 6 | 100.0 | 12 | 2 | US-08-494-301A-16 | Sequence 16, App1 |
| 19 | 6 | 100.0 | 12 | 2 | US-08-480-020B-18 | Sequence 18, App1 |
| 20 | 6 | 100.0 | 12 | 2 | US-08-480-020B-21 | Sequence 21, App1 |
| 21 | 6 | 100.0 | 12 | 2 | US-08-910-618-18 | Sequence 18, App1 |
| 22 | 6 | 100.0 | 12 | 2 | US-08-910-618-21 | Sequence 21, App1 |
| 23 | 6 | 100.0 | 12 | 3 | US-08-779-355-5 | Sequence 5, App1 |
| 24 | 6 | 100.0 | 12 | 3 | US-08-671-824-19 | Sequence 19, App1 |
| 25 | 6 | 100.0 | 12 | 3 | US-08-874-825-88 | Sequence 88, App1 |
| 26 | 6 | 100.0 | 12 | 3 | US-08-938-835A-5 | Sequence 5, App1 |
| 27 | 6 | 100.0 | 12 | 3 | US-08-663-824-88 | Sequence 88, App1 |

ALIGNMENTS

| | | | | | | |
|----|---|-------|----|---|--------------------|-------------------|
| 28 | 6 | 100.0 | 12 | 4 | US-09-243-335-1 | Sequence 1, App1 |
| 29 | 6 | 100.0 | 12 | 4 | US-08-910-322-18 | Sequence 18, App1 |
| 30 | 6 | 100.0 | 12 | 4 | US-08-910-322-21 | Sequence 21, App1 |
| 31 | 6 | 100.0 | 12 | 4 | US-09-281-418-65 | Sequence 65, App1 |
| 32 | 6 | 100.0 | 12 | 4 | US-08-484-939A-18 | Sequence 18, App1 |
| 33 | 6 | 100.0 | 12 | 4 | US-08-484-939A-21 | Sequence 21, App1 |
| 34 | 6 | 100.0 | 12 | 4 | US-09-043-149-48 | Sequence 48, App1 |
| 35 | 6 | 100.0 | 12 | 4 | US-09-231-303-88 | Sequence 88, App1 |
| 36 | 6 | 100.0 | 13 | 1 | US-08-284-746-5 | Sequence 5, App1 |
| 37 | 6 | 100.0 | 13 | 1 | US-08-284-746-16 | Sequence 16, App1 |
| 38 | 6 | 100.0 | 13 | 1 | US-08-050-073-162 | Sequence 162, App |
| 39 | 6 | 100.0 | 14 | 1 | US-08-303-004-23 | Sequence 23, App1 |
| 40 | 6 | 100.0 | 14 | 1 | US-08-050-073-99 | Sequence 99, App1 |
| 41 | 6 | 100.0 | 14 | 1 | US-08-073-161 | Sequence 161, App |
| 42 | 6 | 100.0 | 14 | 1 | US-08-192-941-25 | Sequence 25, App1 |
| 43 | 6 | 100.0 | 14 | 1 | US-08-171-718-119 | Sequence 119, App |
| 44 | 6 | 100.0 | 14 | 2 | US-08-173-489C-324 | Sequence 324, App |
| 45 | 6 | 100.0 | 14 | 2 | US-08-232-087A-6 | Sequence 6, App1 |

RESULT 1
US-08-859-954-49
Sequence 49, Application US/08859954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramin
TITLE OF INVENTION: Design and Optimized Primer Library for
NUMBER OF SEQUENCES: 566
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
City: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,954
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-59900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
HYPOTHETICAL: YES
ANTI-SENSE: YES
US-08-859-954-49
Query Match 100.0%; Score 6; DB 3; Length 8;

Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
11111:
Db 3 GCCTCT 8

RESULT 2
US-08-388-353-459/C
Sequence 459, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-459.

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
11111:
Db 10 GCCTCT 5

RESULT 3
US-08-388-353-460/C
Sequence 460, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
11111:
Db 9 GCCTCT 4

RESULT 4
US-08-388-353-461/C
Sequence 461, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 5
US-08-388-353-462/c
Sequence 462, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 7 GCCTCT 2

RESULT 6
US-08-388-353-463/c
Sequence 463, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 6 GCCTCT 1

RESULT 7
US-08-488-551B-459/c
Sequence 459, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGICLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 10 GCCTCT 5

RESULT 8
US-08-488-551B-460/C
Sequence 460 Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGICLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 9 GCCTCT 4

RESULT 9
US-08-488-551B-461/C
Sequence 461 Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGICLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 10
US-08-488-551B-462/c
Sequence 462, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
|||||
DB 7 GCCTCT 2

RESULT 11
US-08-488-551B-463/c
Sequence 463, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 6 GCCTCT 1

RESULT 12
US-09-235-614-2/c
Sequence 2, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.

APPLICANT: CLARK, CHRISTOPHER L.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-2

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 6 GCCTCT 1

RESULT 13
US-09-235-614-3/C
Sequence 3, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-3

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 7 GCCTCT 2

RESULT 14
US-09-235-614-4/C
Sequence 4, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING

FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 15
US-09-235-614-5/C
Sequence 5, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-5

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 9 GCCTCT 4

Search completed: May 21, 2003, 06:28:35
Job time : 36.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:39:53 ; Search time 238 Seconds
(without alignments)
33.289 Million cell updates/sec

Title: US-09-936-146-2

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues 1657494

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| C 1 | 6 | 100.0 | 9 | US-09-990-186-588 | Sequence 588, App |
| C 2 | 6 | 100.0 | 9 | US-09-990-186-2084 | Sequence 2084, Ap |
| C 3 | 6 | 100.0 | 9 | US-09-990-186-2086 | Sequence 2086, Ap |
| C 4 | 6 | 100.0 | 9 | US-09-990-186-2219 | Sequence 2219, Ap |
| C 5 | 6 | 100.0 | 9 | US-09-990-186-2280 | Sequence 2280, Ap |
| C 6 | 6 | 100.0 | 9 | US-09-990-186-2471 | Sequence 2471, Ap |
| C 7 | 6 | 100.0 | 9 | US-09-990-186-2473 | Sequence 2473, Ap |
| C 8 | 6 | 100.0 | 9 | US-09-989-789-588 | Sequence 588, App |
| C 9 | 6 | 100.0 | 9 | US-09-989-789-2084 | Sequence 2084, Ap |
| C 10 | 6 | 100.0 | 9 | US-09-989-789-2086 | Sequence 2086, Ap |
| C 11 | 6 | 100.0 | 9 | US-09-989-789-2219 | Sequence 2219, Ap |
| C 12 | 6 | 100.0 | 9 | US-09-989-789-2280 | Sequence 2280, Ap |
| C 13 | 6 | 100.0 | 9 | US-09-989-789-2471 | Sequence 2471, Ap |
| C 14 | 6 | 100.0 | 9 | US-09-989-789-2473 | Sequence 2473, Ap |
| C 15 | 6 | 100.0 | 10 | US-09-990-186-617 | Sequence 617, App |
| C 16 | 6 | 100.0 | 10 | US-09-990-186-1279 | Sequence 1279, Ap |
| C 17 | 6 | 100.0 | 10 | US-09-990-186-1308 | Sequence 1308, Ap |
| C 18 | 6 | 100.0 | 10 | US-09-990-186-1313 | Sequence 1313, Ap |
| C 19 | 6 | 100.0 | 10 | US-09-989-789-617 | Sequence 617, App |

| | | | | | |
|------|---|-------|----|---------------------|-------------------|
| C 20 | 6 | 100.0 | 10 | US-09-989-789-1279 | Sequence 1279, Ap |
| C 21 | 6 | 100.0 | 10 | US-09-989-789-1308 | Sequence 1308, Ap |
| C 22 | 6 | 100.0 | 10 | US-09-989-789-1313 | Sequence 1313, Ap |
| C 23 | 6 | 100.0 | 10 | US-10-033-145-153 | Sequence 153, App |
| C 24 | 6 | 100.0 | 10 | US-10-033-145-174 | Sequence 174, App |
| C 25 | 6 | 100.0 | 10 | US-10-033-145-355 | Sequence 355, App |
| C 26 | 6 | 100.0 | 10 | US-10-033-145-566 | Sequence 566, App |
| C 27 | 6 | 100.0 | 10 | US-10-033-145-968 | Sequence 968, App |
| C 28 | 6 | 100.0 | 10 | US-10-033-145-1007 | Sequence 1007, Ap |
| C 29 | 6 | 100.0 | 10 | US-10-033-145-1120 | Sequence 1120, Ap |
| C 30 | 6 | 100.0 | 10 | US-10-033-145-1286 | Sequence 1286, Ap |
| C 31 | 6 | 100.0 | 10 | US-10-033-145-1609 | Sequence 1609, Ap |
| C 32 | 6 | 100.0 | 11 | US-09-249-155-25 | Sequence 25, Appl |
| C 33 | 6 | 100.0 | 11 | US-09-249-155-83 | Sequence 83, Appl |
| C 34 | 6 | 100.0 | 11 | US-09-249-155-207 | Sequence 207, App |
| C 35 | 6 | 100.0 | 11 | US-09-249-155-239 | Sequence 239, App |
| C 36 | 6 | 100.0 | 12 | US-10-131-591A-73 | Sequence 73, Appl |
| C 37 | 6 | 100.0 | 12 | US-09-761-116-1 | Sequence 1, Appl1 |
| C 38 | 6 | 100.0 | 12 | US-09-384-472-18 | Sequence 18, Appl |
| C 39 | 6 | 100.0 | 12 | US-09-384-472-21 | Sequence 21, Appl |
| C 40 | 6 | 100.0 | 14 | US-08-591-4868-102 | Sequence 102, App |
| C 41 | 6 | 100.0 | 14 | US-09-978-600-198 | Sequence 198, App |
| C 42 | 6 | 100.0 | 14 | US-09-998-027-26 | Sequence 26, Appl |
| C 43 | 6 | 100.0 | 14 | US-09-504-231A-1431 | Sequence 1431, Ap |
| C 44 | 6 | 100.0 | 14 | US-09-274-553D-1431 | Sequence 1431, Ap |
| C 45 | 6 | 100.0 | 15 | US-10-056-414-119 | Sequence 119, App |

ALIGNMENTS

RESULT 1

US-09-990-186-588/C
; Sequence 588, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LID, Olang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-588

Query Match

Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY

1 GCCTCU 6

|||||

Db 8 GCCTCT 3

RESULT 2

US-09-990-186-2084/C
; Sequence 2084, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LID, Olang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20

NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2084
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2084

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 3
US-09-990-186-2086/c
Sequence 2086, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2086
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2086

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 4
US-09-990-186-2219/c
Sequence 2219, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2219
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2219

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 5
US-09-990-186-2280/c
Sequence 2280, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2280
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2280

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 6
US-09-990-186-2471/c
Sequence 2471, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2471
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2471

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
DB 8 GCCTCT 3

RESULT 7
US-09-990-186-2473/c

```
; Sequence 2473, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-2473
```

```
Query Match          100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 8
US-09-989-789-588/c
; Sequence 588, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-588
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 9
US-09-989-789-2084/c
; Sequence 2084, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 2084
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2084
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 10
US-09-989-789-2086/c
; Sequence 2086, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2086
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2086
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3
```

```
RESULT 11
US-09-989-789-2219/c
; Sequence 2219, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2219
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2219
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 12
US-09-989-789-2280/c
; Sequence 2280, Application US/09989789
; Patent No. US2002006379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2280
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2280

Query Match 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 13
US-09-989-789-2471/c
; Sequence 2471, Application US/09989789
; Patent No. US2002006379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2471
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2471

Query Match 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 14
US-09-989-789-2473/c
; Sequence 2473, Application US/09989789
; Patent No. US2002006379A1

; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2473

Query Match 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 15
US-09-990-186-617/c
; Sequence 617, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-617

Query Match 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

Search completed: May 21, 2003, 06:44:43
Job time : 238 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 05:13:24 ; Search time 311.5 Seconds

(without alignments)
526.748 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6

Sequence: 1 gccucu 6

Scoring table:

Gapop 60.0 , Gapect 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| C 1 | 100.0 | 10 | 6 | AR128987 | AR128987 Sequence |
| C 2 | 100.0 | 10 | 6 | AR128988 | AR128988 Sequence |
| C 3 | 100.0 | 10 | 6 | AR128989 | AR128989 Sequence |
| C 4 | 100.0 | 10 | 6 | AR128990 | AR128990 Sequence |
| C 5 | 100.0 | 10 | 6 | AR128991 | AR128991 Sequence |
| C 6 | 100.0 | 10 | 6 | AX152349 | AX152349 Sequence |
| C 7 | 100.0 | 10 | 6 | AX152410 | AX152410 Sequence |
| C 8 | 100.0 | 10 | 6 | AX152849 | AX152849 Sequence |
| C 9 | 100.0 | 10 | 6 | AX152886 | AX152886 Sequence |
| C 10 | 100.0 | 10 | 6 | AX153384 | AX153384 Sequence |
| C 11 | 100.0 | 10 | 6 | AX153513 | AX153513 Sequence |
| C 12 | 100.0 | 10 | 6 | E16890 | E16890 DNA sequence |
| C 13 | 100.0 | 11 | 6 | A91501 | A91501 Sequence 28 |
| C 14 | 100.0 | 11 | 6 | A91503 | A91503 Sequence 30 |
| C 15 | 100.0 | 11 | 6 | AX470495 | AX470495 Sequence |
| C 16 | 100.0 | 11 | 6 | AX470740 | AX470740 Sequence |
| C 17 | 100.0 | 11 | 6 | AX470747 | AX470747 Sequence |
| C 18 | 100.0 | 11 | 6 | AX471502 | AX471502 Sequence |
| C 19 | 100.0 | 11 | 6 | AX471630 | AX471630 Sequence |
| C 20 | 100.0 | 11 | 6 | AX471677 | AX471677 Sequence |
| C 21 | 100.0 | 11 | 6 | AX471805 | AX471805 Sequence |
| C 22 | 100.0 | 12 | 6 | A06058 | A06058 Synthetic P |
| C 23 | 100.0 | 12 | 6 | A06059 | A06059 Synthetic P |
| C 24 | 100.0 | 12 | 6 | A15123 | A15123 Nucleotide |
| C 25 | 100.0 | 12 | 6 | A16601 | A16601 Nucleotide |
| C 26 | 100.0 | 12 | 6 | A16602 | A16602 Nucleotide |
| C 27 | 100.0 | 12 | 6 | A47656 | A47656 Sequence 16 |
| C 28 | 100.0 | 12 | 6 | A61520 | A61520 Sequence 89 |
| C 29 | 100.0 | 12 | 6 | A91489 | A91489 Sequence 16 |
| C 30 | 100.0 | 12 | 6 | AR027874 | AR027874 Sequence |
| C 31 | 100.0 | 12 | 6 | AR075465 | AR075465 Sequence |
| C 32 | 100.0 | 12 | 6 | AR075468 | AR075468 Sequence |
| C 33 | 100.0 | 12 | 6 | AR101000 | AR101000 Sequence |
| C 34 | 100.0 | 12 | 6 | AR137925 | AR137925 Sequence |
| C 35 | 100.0 | 12 | 6 | AR153916 | AR153916 Sequence |
| C 36 | 100.0 | 12 | 6 | AR153919 | AR153919 Sequence |
| C 37 | 100.0 | 12 | 6 | AR167701 | AR167701 Sequence |
| C 38 | 100.0 | 12 | 6 | AR178533 | AR178533 Sequence |
| C 39 | 100.0 | 12 | 6 | AR178536 | AR178536 Sequence |
| C 40 | 100.0 | 12 | 6 | AR199100 | AR199100 Sequence |
| C 41 | 100.0 | 12 | 6 | AX233626 | AX233626 Sequence |
| C 42 | 100.0 | 12 | 6 | E29585 | E29585 Method for |
| C 43 | 100.0 | 12 | 6 | E38691 | E38691 Method for |
| C 44 | 100.0 | 12 | 6 | E64117 | E64117 Method for |
| C 45 | 100.0 | 12 | 6 | I07918 | I07918 Sequence 30 |

ALIGNMENTS

RESULT 1

AR128987/c

LOCUS AR128987 10 bp DNA Linear PAT 16-MAY-2001

DEFINITION Sequence 2 from patent US 6183966.

ACCESSION AR128987

VERSION AR128987.1 GI:14116649

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 10)

AUTHORS Gray, D.M. and Clark, C.L.

TITLE Apparatus and method for selectively ranking sequences for antisense targeting

JOURNAL Patent: US 6183966-A 2 06-FEB-2001;

FEATURES Location/Qualifiers
source 1..10
BASE COUNT 2 a 3 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0;

QY 1 GCCUCU 6
DB 6 GCCTCT 1

RESULT 2
LOCUS AR128988 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6183966.
ACCESSION AR128988
VERSION AR128988.1 GI:1411650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURNAL Patent: US 6183966-A 3 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10
BASE COUNT 2 a 3 c 5 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0;

QY 1 GCCUCU 6
DB 7 GCCTCT 2

RESULT 3
LOCUS AR128989 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 4 from patent US 6183966.
ACCESSION AR128989
VERSION AR128989.1 GI:1411651
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURNAL Patent: US 6183966-A 4 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10
BASE COUNT 3 a 3 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0;

QY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 4
LOCUS AR128990 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6183966.
ACCESSION AR128990
VERSION AR128990.1 GI:1411652
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURNAL Patent: US 6183966-A 5 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10
BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0;

QY 1 GCCUCU 6
DB 9 GCCTCT 4

RESULT 5
LOCUS AR128991 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6183966.
ACCESSION AR128991
VERSION AR128991.1 GI:1411653
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURNAL Patent: US 6183966-A 6 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10
BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0;

QY 1 GCCUCU 6
DB 10 GCCTCT 5

RESULT 6
LOCUS AX152349 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 264 from Patent WO0138577.
ACCESSION AX152349
VERSION AX152349.1 GI:14534000
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 264 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
    source
        1..10
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT 0 a 5 c 2 g 3 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
Db 1 GCCTCT 6

RESULT 7
AX152410 10 bp DNA PAT 22-JUN-2001
LOCUS Sequence 325 from Patent WO0138577.
DEFINITION AX152410.
ACCESSION AX152410.
VERSION AX152410.1 GI:14534061
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 325 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
    source
        1..10
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT 1 a 5 c 2 g 2 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
Db 1 GCCTCT 6

RESULT 8
AX152849 10 bp DNA PAT 22-JUN-2001
LOCUS Sequence 764 from Patent WO0138577.
DEFINITION AX152849
ACCESSION AX152849
VERSION AX152849.1 GI:14534500
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 764 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
    source
        1..10
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
Db 1 GCCTCT 3

RESULT 9
AX152886 10 bp DNA PAT 22-JUN-2001
LOCUS Sequence 801 from Patent WO0138577.
DEFINITION AX152886
ACCESSION AX152886
VERSION AX152886.1 GI:14534537
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 801 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
    source
        1..10
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT 5 a 1 c 3 g 1 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
Db 8 GCCTCT 3

RESULT 10
AX153384 10 bp DNA PAT 22-JUN-2001
LOCUS Sequence 1299 from Patent WO0138577.
DEFINITION AX153384
ACCESSION AX153384
VERSION AX153384.1 GI:14535035
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 1299 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
    source
        1..10
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|

```

Db 2 GCCTCT 7.

RESULT 11
AX153513 10 bp DNA Linear PAT 22-JUN-2001
DEFINITION Sequence 1428 from Patent WO0138577.
ACCESSION AX153513
VERSION AX153513.1 GI:14535164
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Velculescu, V.E., Vogelstein, B. and Kinzler, K.W.
TITLE Human transposomes
JOURNAL Patent: WO 0138577-A 1428 31-MAY-2001;
The Johns Hopkins University (US)
Location/Qualifiers

FEATURES
source 1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 0 a 4 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:|
Db 1 GCCTCT 6

RESULT 12
E16890 10 bp DNA Linear PAT 28-JUL-1999
LOCUS E16890
DEFINITION DNA sequence required for efficient protein transcription in
Brevibacterium flavum.
ACCESSION E16890
VERSION E16890.1 GI:5711573
KEYWORDS JP 1998229881-A/31.
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 (bases 1 to 10)
AUTHORS Kobayashi, M., Man, T. and Yugawa, H.
TITLE DNA HAVING SEQUENCE CAPABLE OF EFFICIENTLY TRANSLATING PROTEIN IN
CORYNEFORM BACTERIA
JOURNAL Patent: JP 1998229881-A 31 02-SEP-1998;
MITSUBISHI CHEM CORP
OS Brevibacterium flavum
PN JP 1998229881-A/31
PD 02-SEP-1998
PF 19-FEB-1997 JP 1997035338
PI KOBAYASHI MIKI, MAN TOMOKO, YUGAWA HIDEAKI
PC C12N15/09.C07H21/04.C12N1/21//C12N9/38.C12Q1/68.(C12N15/09, PC
C12R1.19)
PC (C12N1/21,C12N1.13),(C12N9/38,C12N1.19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1. .10
/strain="MJ-233",
/organism="Brevibacterium flavum" FT
/location/Qualifiers
1. .10
/organism="Corynebacterium glutamicum"

FEATURES
source

BASE COUNT 3 a 1 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.8e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:|
Db 9 GCCTCT 4

RESULT 13
A91501 11 bp DNA Linear PAT 22-JAN-2000
LOCUS A91501
DEFINITION Sequence 28 from Patent WO9824928.
ACCESSION A91501
VERSION A91501.1 GI:6740456
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 11)
AUTHORS Pallisgaard, N. and Hokland, P.
TITLE DETECTION OF CHROMOSOMAL ABNORMALITIES
JOURNAL Patent: WO 9824928-A 28 11-JUN-1998;
PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers

FEATURES
source 1. .11
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 1 a 5 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 66.7%; Pred. No. 2.7e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:|
Db 5 GCCTCT 10

RESULT 14
A91503 11 bp DNA Linear PAT 22-JAN-2000
LOCUS A91503
DEFINITION Sequence 30 from Patent WO9824928.
ACCESSION A91503
VERSION A91503.1 GI:6740458
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 11)
AUTHORS Pallisgaard, N. and Hokland, P.
TITLE DETECTION OF CHROMOSOMAL ABNORMALITIES
JOURNAL Patent: WO 9824928-A 30 11-JUN-1998;
PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers

FEATURES
source 1. .11
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 0 a 5 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 66.7%; Pred. No. 2.7e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:|
Db 5 GCCTCT 10

```

RESULT 15
LOCUS AX470495/c 11 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 72 from Patent WO02053773.
ACCESSION AX470495
VERSION AX470495.1 GI:22205620
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hofmann,K., Conradt,M. and Petersohn,D.
TITLE Method for determining skin stress or skin aging in vitro
JOURNAL Patent: WO 02053773-A 72 11-JUL-2002;
HENKEL KGNA (DE)
FEATURES
source Location/Qualifiers
1..11
BASE COUNT 6 a 2 c 3 g 0 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 66.7%; Pred. No. 2,7e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
|||:|:
DB 8 GCCCTC 3

```

Search completed: May 21, 2003, 07:17:19
 Job time : 332.5 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:50:53 ; Search time 148.25 Seconds

(without alignments)
91.143 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccuuc 6

Scoring table: OLIGO_MNC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N.Geneseq_101002.*

```
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 6 | 100.0 | 8 | 20 | AAAX29578 |
| 2 | 6 | 100.0 | 8 | 21 | AAA80736 |
| 3 | 6 | 100.0 | 9 | 24 | ABO71469 |
| 4 | 6 | 100.0 | 9 | 24 | ABO71786 |
| 5 | 6 | 100.0 | 9 | 24 | ABO71788 |
| 6 | 6 | 100.0 | 9 | 24 | ABO71921 |
| 7 | 6 | 100.0 | 9 | 24 | ABO71982 |
| 8 | 6 | 100.0 | 9 | 24 | ABO72173 |
| 9 | 6 | 100.0 | 9 | 24 | ABO72175 |

| | | | | | | | |
|---|----|---|-------|----|----|----------|--------------------|
| C | 10 | 6 | 100.0 | 10 | 16 | AAQ96863 | HIV-1 NL4-3 nef ge |
| C | 11 | 6 | 100.0 | 10 | 16 | AAQ96864 | HIV-1 NL4-3 nef ge |
| C | 12 | 6 | 100.0 | 10 | 16 | AAQ96865 | HIV-1 NL4-3 nef ge |
| C | 13 | 6 | 100.0 | 10 | 16 | AAQ96866 | HIV-1 NL4-3 nef ge |
| C | 14 | 6 | 100.0 | 10 | 16 | AAQ96867 | HIV-1 NL4-3 nef ge |
| C | 15 | 6 | 100.0 | 10 | 20 | AAH86201 | SAGE tag used to 1 |
| C | 16 | 6 | 100.0 | 10 | 20 | AAH86201 | Antisense oligonuc |
| C | 17 | 6 | 100.0 | 10 | 20 | AAH86201 | Antisense oligonuc |
| C | 18 | 6 | 100.0 | 10 | 20 | AAH86201 | Antisense oligonuc |
| C | 19 | 6 | 100.0 | 10 | 20 | AAH86201 | Antisense oligonuc |
| C | 20 | 6 | 100.0 | 10 | 20 | AAH86201 | Antisense oligonuc |
| C | 21 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 22 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 23 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 24 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 25 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 26 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 27 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 28 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 29 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 30 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 31 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 32 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 33 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 34 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 35 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 36 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 37 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 38 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 39 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 40 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 41 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 42 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 43 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 44 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |
| C | 45 | 6 | 100.0 | 10 | 21 | AAH86201 | Antisense oligonuc |

ALIGNMENTS

| | | |
|----------|---|----------------------|
| RESULT 1 | AAAX29578 | standard; DNA; 8 Bp. |
| ID | AAAX29578; | |
| AC | AAAX29578; | |
| XX | 03-JUN-1999 | (first entry) |
| DT | 03-JUN-1999 | (first entry) |
| DE | Primer for human G-protein coupled receptor genes. | |
| XX | Nucleic acid amplification; nuclear receptor; G-protein coupled receptor; | |
| KW | apoptosis; DNA repair; DNA replication; plant biology; agriculture; | |
| KW | human; veterinary medicine; reproduction; microbiology; hybridization; | |
| KW | environmental science; DNA fingerprinting; PCR primer; ss. | |
| XX | Synthetic. | |
| OS | Homo sapiens. | |
| XX | WO911823-A2. | |
| PN | 11-MAR-1999. | |
| PD | 04-SEP-1998; | 98WO-US18392. |
| PF | 05-SEP-1997; | 97US-0925816. |
| XX | (KIMM-) KIMMEL CANCER CENT SIDNEY. | |
| PA | McClelland M, Pesole G; | |
| PI | WPI; 1999-205200/17. | |
| XX | | |
| DR | | |
| XX | | |

PT Subset of primers able to amplify group of related sequences
 XX
 PS Claim 21; Page 75; 92pp; English.
 XX
 CC The invention provides primers (AA29501-X29679) for identifying
 CC sequences encoding structurally or functionally related proteins such as
 CC nuclear or G-protein coupled receptors, apoptosis-related or DNA
 CC repair/replication proteins. The identified sequences are broadly useful
 CC in plant biology, agriculture, human or veterinary medicine,
 CC reproduction, microbiology or environmental science, e.g. to study
 CC expression of nuclear receptors at different stages of tissue development
 CC or after treatment with particular drugs. It is also used for DNA
 CC fingerprinting (to generate products useful for differential
 CC hybridisation), or, where a 3'-anchor primer is used, to isolate the
 CC 3'-ends of mRNA sequences. Sequences AA29526-X29583 represent claimed
 CC primers specific for human G-protein coupled receptor genes.
 XX
 SQ Sequence 8 BP; 0 A; 4 C; 2 G; 2 T; 0 other;

Query Match 100.0%; Score 6; DB 20; Length 8;
 Best Local Similarity 66.7%; Pred. No. 2.6e+08;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 |||:
 Db 1 GCCCTCT 6

RESULT 2
 AAA80736
 ID AAA80736 standard; DNA; 8 BP.
 XX
 AC AAA80736;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE A. thaliana primer walking octamer SEQ ID NO: 49.
 XX
 KW Primer walking; octamer; primer; DNA sequencing; PCR; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US6083695-A.
 XX
 PD 04-JUL-2000.
 XX
 PF 21-MAY-1997; 97US-0859954.
 XX
 PR 15-APR-1996; 96US-0632782.
 XX
 PA (UYHO-) UNIV HOUSTON.
 XX
 PA (HARD/) HARDIN S H.
 XX
 PI Hardin PE, Hardin SH, Homayouni R;
 XX
 DR WPI; 2000-474852/41.
 XX
 PT Sequencing an unknown DNA molecule for the polymerase chain reaction
 PT and other primer processes comprises primer walking of octamer
 PT oligonucleotides -
 XX
 PS Example 8; Column 49-50; 161pp; English.
 XX
 CC This invention describes a novel method for sequencing an unknown DNA
 CC molecule which comprises selecting a library primer from an octamer
 CC oligonucleotide library consisting of 48 8-bp sequences and
 CC corresponding complementary sequences, where the library primer is
 CC complementary to a known sequence adjacent to the unknown sequence or
 CC is complementary to a sequence in a known extension product. The method
 CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
 CC which make use of primers. The octamers are used to identify coding
 CC sequences. Primer walking using the octamer libraries is advantageous
 CC over other sequencing methods because it does not require multiple

CC Cloning steps nor subsequent template preparations, and it is a
 CC directed and methodical approach. AAA80688-A81253 represent the octamer
 CC primers used in the primer walking method of the invention.
 XX
 SQ Sequence 8 BP; 0 A; 3 C; 1 G; 4 T; 0 other;

Query Match 100.0%; Score 6; DB 21; Length 8;
 Best Local Similarity 66.7%; Pred. No. 2.6e+08;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 |||:
 Db 3 GCCCTCT 8

RESULT 3
 ABQ71469/C
 ID ABQ71469 standard; DNA; 9 BP.
 XX
 AC ABQ71469;
 XX
 DT 28-AUG-2002 (first entry)
 XX
 DE Zinc finger protein related oligonucleotide target SEQ ID NO:588.
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
 XX
 OS Homo sapiens.
 XX
 OS Synthetic.
 XX
 PN WO200242459-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 20-NOV-2001; 2001WO-US43438.
 XX
 PR 20-NOV-2000; 2000US-0716637.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Liu Q;
 XX
 DR WPI; 2002-500284/53.
 XX
 PT New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering.
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 XX
 PS Example 1; Page 44; 81pp; English.
 XX
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (W) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and AB448191 to ABP1230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;
SQ Best Local Similarity 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.3e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 4
ABQ71786/c
ID ABQ71786 standard; DNA; 9 BP.
XX ABQ71786;
XX 28-AUG-2002 (first entry)
XX Zinc finger protein related oligonucleotide target SEQ ID NO:2084.
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX Homo sapiens.
XX Synthetic.
XX MO200242459-A2.
XX 30-MAY-2002.
XX 20-NOV-2001; 2001MO-US43438.
XX 20-NOV-2000; 2000US-0716637.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX Liu Q;
XX WPI; 2002-500284/53.
XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering,
XX comprises first, second and third zinc fingers, ordered from N- to
XX C-terminus -
XX Example 1; Page 55; 81pp; English.
XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target subsequence. Also described are: (1) a polypeptide
XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target subsequence, selecting the F2 zinc finger such
XX that it binds to the S2 target subsequence, and selecting the F3 zinc
XX finger such that it binds to the S3 target subsequence, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target sites having the nucleotide G in the 5'-most position of the
XX subsequence. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

XX Query Match 100.0%; Score 6; DB 24; Length 9;
XX Best Local Similarity 66.7%; Pred. No. 2.3e+08;
XX Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 5
ABQ71788/c
ID ABQ71788 standard; DNA; 9 BP.
XX ABQ71788;
XX 28-AUG-2002 (first entry)
XX Zinc finger protein related oligonucleotide target SEQ ID NO:2086.
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX Homo sapiens.
XX Synthetic.
XX MO200242459-A2.
XX 30-MAY-2002.
XX 20-NOV-2001; 2001MO-US43438.
XX 20-NOV-2000; 2000US-0716637.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX Liu Q;
XX WPI; 2002-500284/53.
XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering,
XX comprises first, second and third zinc fingers, ordered from N- to
XX C-terminus -
XX Example 1; Page 55; 81pp; English.
XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target subsequence. Also described are: (1) a polypeptide
XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target subsequence, selecting the F2 zinc finger such
XX that it binds to the S2 target subsequence, and selecting the F3 zinc
XX finger such that it binds to the S3 target subsequence, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target sites having the nucleotide G in the 5'-most position of the
XX subsequence. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

OY 1 GCCUCU 6
||||:|
DB 8 GCCTCT 3

RESULT 6

ABQ71921/c
ID ABQ71921 standard; DNA; 9 BP.

XX AC ABQ71921;

DT 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2219.

XX KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.
XX OS Synthetic.

PN WO200242459-A2.

XX PD 30-MAY-2002.

XX PF 20-NOV-2001; 2001WO-US43438.

XX PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
PF gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus

XX Example 1; Page 58; 81pp; English.

PS The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsequence. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsequence, selecting the F2 zinc finger such
CC that it binds to the S2 target subsequence, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsequence, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsequences having the nucleotide G in the 5'-most position of the
CC subsequence. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject. In diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

SO Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 66.7%; Pred. No. 2.3e+08;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:|

DB 8 GCCTCT 3

RESULT 7

ABQ71982/c
ID ABQ71982 standard; DNA; 9 BP.

XX AC ABQ71982;

DT 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2280.

XX KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.
XX OS Synthetic.

PN WO200242459-A2.

XX PD 30-MAY-2002.

XX PF 20-NOV-2001; 2001WO-US43438.

XX PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
PF gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus

XX Example 1; Page 59; 81pp; English.

PS The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsequence. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsequence, selecting the F2 zinc finger such
CC that it binds to the S2 target subsequence, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsequence, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsequences having the nucleotide G in the 5'-most position of the
CC subsequence. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject. In diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

SO Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 66.7%; Pred. No. 2.3e+08;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:|

DB 8 GCCTCT 3

```

RESULT 8
ABQ72173/c
ID ABQ72173 standard; DNA: 9 BP.
XX
AC ABQ72173;
XX
DT 28-AUG-2002 (first entry)
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2471.
XX
KM Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200242459-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Liu Q;
XX
DR WPI; 2002-500284/53.
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus
XX
PS Example 1; Page 63; 81pp; English.
XX
CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.3e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCUCU 6
|||:|
DB 8 GCCTCT 3

```

```

XX
AC ABQ72175;
XX
DT 28-AUG-2002 (first entry)
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2473.
XX
KM Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200242459-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PR 20-NOV-2000; 2000US-0716637.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Liu Q;
XX
DR WPI; 2002-500284/53.
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus
XX
PS Example 1; Page 63; 81pp; English.
XX
CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.3e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCUCU 6
|||:|
DB 8 GCCTCT 3

```

```

RESULT 9
ABQ72175/c
ID ABQ72175 standard; DNA: 9 BP.

```

```

RESULT 10
AAQ96863/c
ID AAQ96863 standard; DNA: 10 BP.
XX
AC AAQ96863;
XX

```

DT 26-MAR-1996 (first entry)
 XX
 DE HIV-1 NL4-3 nef gene nucleotide deletion 458.
 XX
 KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PE 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.
 XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 CC (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX
 DR WPI: 1995-293115/38.
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX
 PS Claim 13; Page 194; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
 CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 1 C; 4 G; 1 T; 0 other;
 XX
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GCCUCU 6
 |||:|:
 Db 10 GCCTCT 5

RESULT 11
 AA096864/c
 ID AA096864 standard; DNA: 10 BP.
 XX
 AC AA096864;
 XX
 DT 26-MAR-1996 (first entry)
 XX
 DE HIV-1 NL4-3 nef gene nucleotide deletion 459.
 XX
 KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PE 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.

XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX
 DR WPI: 1995-293115/38.
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX
 PS Claim 13; Page 194; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
 CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;
 XX
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GCCUCU 6
 |||:|:
 Db 9 GCCTCT 4

RESULT 12
 AA096865/c
 ID AA096865 standard; DNA: 10 BP.
 XX
 AC AA096865;
 XX
 DT 26-MAR-1996 (first entry)
 XX
 DE HIV-1 NL4-3 nef gene nucleotide deletion 460.
 XX
 KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PE 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.
 XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 CC (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX
 DR WPI: 1995-293115/38.
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX
 PS Claim 13; Page 194; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
 CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX

SQ Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
 |||:|:
 Db 8 GCCTCT 3

RESULT 13
 AA096866/c
 ID AA096866 standard; DNA: 10 BP.

XX
 AC AA096866;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 461.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX 17-AUG-1995:

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX WPI; 1995-293115/38.

DR New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 XX or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX

PS Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decaunucleotides (AA096406-097018) from the nef gene and/or
 CC 1 or more decaunucleotides (AA097019-097166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX

SQ Sequence 10 BP; 5 A; 2 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
 |||:|:
 Db 7 GCCTCT 2

RESULT 14

AA096867/c
 ID AA096867 standard; DNA: 10 BP.

XX
 AC AA096867;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 462.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX WPI; 1995-293115/38.

DR New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 XX or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 XX

PS Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
 CC or more decaunucleotides (AA096406-097018) from the nef gene and/or
 CC 1 or more decaunucleotides (AA097019-097166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX

SQ Sequence 10 BP; 4 A; 2 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
 |||:|:
 Db 6 GCCTCT 1

RESULT 15

AA096201
 ID AA096201 standard; DNA: 10 BP.

XX
 AC AA096201;

DT 22-SEP-1999 (first entry)

DE SAGE tag used to identify transcripts which are enhanced by p53.

XX p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity;

XX neoplastic; p53 binding site; PIG-3 promoter; SAGE tag; ss.

OS Homo sapiens.

XX WO9914356-A2.

XX 25-MAR-1999.

XX

XX

XX

```

XX 17-SEP-1998; 98WD-US19300.
PF 30-MAR-1998; 98US-0079817.
XX 17-SEP-1997; 97US-0059153.
PR 17-SEP-1997; 97US-0059153.
XX
PA (UYJO ) UNIV JOHNS HOPKINS.
XX
PI Klinzler KW, Polyak K, Vogelstein B;
XX WPI; 1999-443793/37.
DR
XX Use of p53 transcription tags to determine p53 status in, e.g.
PT cancer diagnosis
XX
PS Example 1; Page 25; 73pp; English.
XX
CC The specification describes the use of p53 transcription tags for
CC developing products to determine p53 status, to diagnose cancer
CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
CC A method for diagnosing cancer or determining p53 status in a sample
CC suspected for being neoplastic comprises comparing the level of
CC transcription of an RNA transcript in a first sample (s1) of a first
CC tissue (t1) to the level of transcription of the transcript in a second
CC sample (s2) of a second tissue (s2), where s1 is suspected of being
CC neoplastic and s2 is a normal human tissue (of the same type) and the
CC transcript is identified by a tag; and categorizing s1 as neoplastic
CC or as having a mutant p53 when transcription is found to be the same
CC or lower in the first, than in s2. The methods and products can be used
CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
CC or carcinogenicity of a test agent. AAX86201-33 represent SAGE tags
CC used to identify transcripts which are enhanced by p53.
XX
SQ Sequence 10 BP; 0 A; 6 C; 1 G; 3 T; 0 other;

```

Query Match 100.0%; Score 6; DB 20; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.9e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GCCUCU 6
   |||:|:
Db 4 GCCTCT 9

```

Search completed: May 21, 2003, 06:54:55
 Job time : 148.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:28:43 ; Search time 34.75 Seconds
(without alignments)
52.951 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccuccu 6

Scoring table: OLIGO.MDC
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents.NA:*
1: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
3: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*
5: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*
6: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 6 | 100.0 | 8 | 3 | US-08-859-954-49 |
| 2 | 6 | 100.0 | 10 | 3 | US-08-388-353-459 |
| 3 | 6 | 100.0 | 10 | 3 | US-08-388-353-460 |
| 4 | 6 | 100.0 | 10 | 3 | US-08-388-353-461 |
| 5 | 6 | 100.0 | 10 | 3 | US-08-388-353-462 |
| 6 | 6 | 100.0 | 10 | 3 | US-08-388-353-463 |
| 7 | 6 | 100.0 | 10 | 3 | US-08-488-551B-459 |
| 8 | 6 | 100.0 | 10 | 3 | US-08-488-551B-460 |
| 9 | 6 | 100.0 | 10 | 3 | US-08-488-551B-461 |
| 10 | 6 | 100.0 | 10 | 3 | US-08-488-551B-462 |
| 11 | 6 | 100.0 | 10 | 3 | US-08-488-551B-463 |
| 12 | 6 | 100.0 | 10 | 4 | US-09-235-614-2 |
| 13 | 6 | 100.0 | 10 | 4 | US-09-235-614-3 |
| 14 | 6 | 100.0 | 10 | 4 | US-09-235-614-4 |
| 15 | 6 | 100.0 | 10 | 4 | US-09-235-614-5 |
| 16 | 6 | 100.0 | 10 | 4 | US-09-235-614-6 |
| 17 | 6 | 100.0 | 10 | 4 | US-09-154-750A-4 |
| 18 | 6 | 100.0 | 12 | 2 | US-08-494-301A-16 |
| 19 | 6 | 100.0 | 12 | 2 | US-08-480-020B-18 |
| 20 | 6 | 100.0 | 12 | 2 | US-08-480-020B-21 |
| 21 | 6 | 100.0 | 12 | 2 | US-08-910-618-18 |
| 22 | 6 | 100.0 | 12 | 2 | US-08-910-618-21 |
| 23 | 6 | 100.0 | 12 | 2 | US-08-779-355-5 |
| 24 | 6 | 100.0 | 12 | 3 | US-08-671-824-19 |
| 25 | 6 | 100.0 | 12 | 3 | US-08-874-825-88 |
| 26 | 6 | 100.0 | 12 | 3 | US-08-938-835A-5 |
| 27 | 6 | 100.0 | 12 | 3 | US-08-663-824-88 |

| | | | | | | |
|----|---|-------|----|---|--------------------|--------------------|
| 28 | 6 | 100.0 | 12 | 4 | US-09-243-335-1 | Sequence 1, Appl |
| 29 | 6 | 100.0 | 12 | 4 | US-08-910-322-18 | Sequence 18, Appl |
| 30 | 6 | 100.0 | 12 | 4 | US-08-910-322-21 | Sequence 21, Appl |
| 31 | 6 | 100.0 | 12 | 4 | US-09-281-418-65 | Sequence 65, Appl |
| 32 | 6 | 100.0 | 12 | 4 | US-08-484-939A-18 | Sequence 18, Appl |
| 33 | 6 | 100.0 | 12 | 4 | US-08-484-939A-21 | Sequence 21, Appl |
| 34 | 6 | 100.0 | 12 | 4 | US-09-043-149-48 | Sequence 48, Appl |
| 35 | 6 | 100.0 | 12 | 4 | US-09-231-303-88 | Sequence 88, Appl |
| 36 | 6 | 100.0 | 13 | 1 | US-08-284-746-5 | Sequence 5, Appl |
| 37 | 6 | 100.0 | 13 | 1 | US-08-284-746-16 | Sequence 16, Appl |
| 38 | 6 | 100.0 | 13 | 1 | US-08-050-073-162 | Sequence 162, Appl |
| 39 | 6 | 100.0 | 14 | 1 | US-08-303-004-23 | Sequence 23, Appl |
| 40 | 6 | 100.0 | 14 | 1 | US-08-050-073-99 | Sequence 99, Appl |
| 41 | 6 | 100.0 | 14 | 1 | US-08-050-073-161 | Sequence 161, Appl |
| 42 | 6 | 100.0 | 14 | 1 | US-08-192-941-25 | Sequence 25, Appl |
| 43 | 6 | 100.0 | 14 | 1 | US-08-171-718-119 | Sequence 119, Appl |
| 44 | 6 | 100.0 | 14 | 2 | US-08-173-489C-324 | Sequence 324, Appl |
| 45 | 6 | 100.0 | 14 | 2 | US-08-232-087A-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1
US-08-859-954-49
Sequence 49, Application US/08859954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramin
APPLICANT: Hardin, Paul E.
TITLE OF INVENTION: Design and Optimized Primer Library for
NUMBER OF SEQUENCES: 566
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,954
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ. ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
HYPOTHEICAL: YES
ANTI-SENSE: YES
US-08-859-954-49

Query Match 100.0%; Score 6; DB 3; Length 8;

Best Local Similarity 66.7%; Pred. No. 3.6e+07;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCUCU 6
| | | | |
DB 3 GCCTCT 8

RESULT 2
US-08-388-353-459/C
; Sequence 459, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 459:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-388-353-459
Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCUCU 6
| | | | |
DB 10 GCCTCT 5

RESULT 3
US-08-388-353-460/C
; Sequence 460, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCUCU 6
| | | | |
DB 9 GCCTCT 4

RESULT 4
US-08-388-353-461/C
; Sequence 461, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|
Db 8 GCCTCT 3

RESULT 5
US-08-388-353-462/c
Sequence 462, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|
Db 7 GCCTCT 2

RESULT 6
US-08-388-353-463/c
Sequence 463, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|
Db 6 GCCTCT 1

RESULT 7
US-08-488-551B-459/c
Sequence 459, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGICLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. NO. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 10 GCCTCT 5

RESULT 8
US-08-488-551B-460/C
Sequence 460, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGICLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. NO. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 9 GCCTCT 4

RESULT 9
US-08-488-551B-461/C
Sequence 461, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGICLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 10
US-08-488-551B-462/c
Sequence 462, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
DB 7 GCCTCT 2

RESULT 11
US-08-488-551B-463/c
Sequence 463, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
DB 6 GCCTCT 1

RESULT 12
US-09-235-614-2/c
Sequence 2, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.

APPLICANT: CLARK, CHRISTOPHER L.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-2

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
Db 6 GCCTCT 1

RESULT 13
US-09-235-614-3/c
Sequence 3, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-3

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
Db 7 GCCTCT 2

RESULT 14
US-09-235-614-4/c
Sequence 4, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 15
US-09-235-614-5/c
Sequence 5, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.
TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
FILE REFERENCE: 91556/66384
CURRENT APPLICATION NUMBER: US/09/235,614
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/808,474
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: 08/320,507
PRIOR FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-5

Query Match 100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.4e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
Db 9 GCCTCT 4

Search completed: May 21, 2003, 08:40:18
Job time : 34.75 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:44:53 ; Search time 65 Seconds
(without alignments)
121.889 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccucu 6

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 828747 seqs, 660231138 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published_Applications_NA:

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1 | 100.0 | 6 | 9 | US-09-990-186-588 | Sequence 588, App |
| 2 | 100.0 | 6 | 9 | US-09-990-186-2084 | Sequence 2084, App |
| 3 | 100.0 | 6 | 9 | US-09-990-186-2086 | Sequence 2086, App |
| 4 | 100.0 | 6 | 9 | US-09-990-186-2219 | Sequence 2219, App |
| 5 | 100.0 | 6 | 9 | US-09-990-186-2280 | Sequence 2280, App |
| 6 | 100.0 | 6 | 9 | US-09-990-186-2471 | Sequence 2471, App |
| 7 | 100.0 | 6 | 9 | US-09-990-186-2473 | Sequence 2473, App |
| 8 | 100.0 | 6 | 9 | US-09-989-789-588 | Sequence 588, App |
| 9 | 100.0 | 6 | 9 | US-09-989-789-2084 | Sequence 2084, App |
| 10 | 100.0 | 6 | 9 | US-09-989-789-2086 | Sequence 2086, App |
| 11 | 100.0 | 6 | 9 | US-09-989-789-2219 | Sequence 2219, App |
| 12 | 100.0 | 6 | 9 | US-09-989-789-2280 | Sequence 2280, App |
| 13 | 100.0 | 6 | 9 | US-09-989-789-2471 | Sequence 2471, App |
| 14 | 100.0 | 6 | 9 | US-09-989-789-2473 | Sequence 2473, App |
| 15 | 100.0 | 6 | 9 | US-09-990-186-617 | Sequence 617, App |
| 16 | 100.0 | 6 | 9 | US-09-990-186-1279 | Sequence 1279, App |
| 17 | 100.0 | 6 | 9 | US-09-990-186-1308 | Sequence 1308, App |
| 18 | 100.0 | 6 | 9 | US-09-990-186-1313 | Sequence 1313, App |
| 19 | 100.0 | 6 | 9 | US-09-989-789-617 | Sequence 617, App |

| | | | | | | |
|----|---|-------|----|----|---------------------|--------------------|
| 20 | 6 | 100.0 | 10 | 10 | US-09-989-789-1279 | Sequence 1279, App |
| 21 | 6 | 100.0 | 10 | 10 | US-09-989-789-1308 | Sequence 1308, App |
| 22 | 6 | 100.0 | 10 | 10 | US-09-989-789-1313 | Sequence 1313, App |
| 23 | 6 | 100.0 | 10 | 12 | US-10-033-145-153 | Sequence 153, App |
| 24 | 6 | 100.0 | 10 | 12 | US-10-033-145-174 | Sequence 174, App |
| 25 | 6 | 100.0 | 10 | 12 | US-10-033-145-355 | Sequence 355, App |
| 26 | 6 | 100.0 | 10 | 12 | US-10-033-145-566 | Sequence 566, App |
| 27 | 6 | 100.0 | 10 | 12 | US-10-033-145-968 | Sequence 968, App |
| 28 | 6 | 100.0 | 10 | 12 | US-10-033-145-1007 | Sequence 1007, App |
| 29 | 6 | 100.0 | 10 | 12 | US-10-033-145-1120 | Sequence 1120, App |
| 30 | 6 | 100.0 | 10 | 12 | US-10-033-145-1286 | Sequence 1286, App |
| 31 | 6 | 100.0 | 10 | 12 | US-10-033-145-1609 | Sequence 1609, App |
| 32 | 6 | 100.0 | 11 | 9 | US-09-249-155-83 | Sequence 83, App |
| 33 | 6 | 100.0 | 11 | 9 | US-09-249-155-207 | Sequence 207, App |
| 34 | 6 | 100.0 | 11 | 9 | US-09-249-155-239 | Sequence 239, App |
| 35 | 6 | 100.0 | 11 | 9 | US-09-761-116-73 | Sequence 73, App |
| 36 | 6 | 100.0 | 12 | 9 | US-09-384-472-18 | Sequence 18, App |
| 37 | 6 | 100.0 | 12 | 10 | US-09-384-472-21 | Sequence 21, App |
| 38 | 6 | 100.0 | 12 | 10 | US-08-591-486B-102 | Sequence 102, App |
| 39 | 6 | 100.0 | 12 | 10 | US-09-978-600-198 | Sequence 198, App |
| 40 | 6 | 100.0 | 14 | 9 | US-09-998-027-26 | Sequence 26, App |
| 41 | 6 | 100.0 | 14 | 9 | US-09-504-231A-1431 | Sequence 1431, App |
| 42 | 6 | 100.0 | 14 | 10 | US-09-274-553D-1431 | Sequence 1431, App |
| 43 | 6 | 100.0 | 14 | 10 | US-10-056-414-119 | Sequence 119, App |
| 44 | 6 | 100.0 | 15 | 9 | | |
| 45 | 6 | 100.0 | 15 | 9 | | |

ALIGNMENTS

```
RESULT 1
US-09-990-186-588/c
; Sequence 588, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-588

Query Match          100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0;

QY      1 GCCUCU 6
      |||::|
Db      8 GCCCTC 3

RESULT 2
US-09-990-186-2084/c
; Sequence 2084, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
```

NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2084
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2084

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 3
US-09-990-186-2086/c
Sequence 2086, Application US/09990186
Publication No. US20030068675a1
GENERAL INFORMATION:

APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2086
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2086

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 4
US-09-990-186-2219/c

Sequence 2219, Application US/09990186
Publication No. US20030068675a1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2219
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2219

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 5
US-09-990-186-2280/c

Sequence 2280, Application US/09990186
Publication No. US20030068675a1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2280
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2280

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 6
US-09-990-186-2471/c

Sequence 2471, Application US/09990186
Publication No. US20030068675a1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2471
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2471

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
DB 8 GCCTCT 3

RESULT 7
US-09-990-186-2473/c

```
; Sequence 2473, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-2473
```

```
Query Match          100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCUCU 6
      |||:|:
Db      8 GCCTCT 3
```

```
RESULT 8
US-09-989-789-588/c
; Sequence 588, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-588
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCUCU 6
      |||:|:
Db      8 GCCTCT 3
```

```
RESULT 9
US-09-989-789-2084/c
; Sequence 2084, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
```

```
; SEQ ID NO 2084
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2084
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCUCU 6
      |||:|:
Db      8 GCCTCT 3
```

```
RESULT 10
US-09-989-789-2086/c
; Sequence 2086, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2086
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2086
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GCCUCU 6
      |||:|:
Db      8 GCCTCT 3
```

```
RESULT 11
US-09-989-789-2219/c
; Sequence 2219, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2219
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2219
```

```
Query Match          100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
```

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCCTCT 3

RESULT 12
US-09-989-789-2280/c
; Sequence 2280, Application US/09989789
; Patent No. US2002006379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2280
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2280

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCCTCT 3

RESULT 13
US-09-989-789-2471/c
; Sequence 2471, Application US/09989789
; Patent No. US2002006379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2471
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2471

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCCTCT 3

RESULT 14
US-09-989-789-2473/c
; Sequence 2473, Application US/09989789
; Patent No. US2002006379A1

GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2473

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCCTCT 3

RESULT 15
US-09-990-186-617/c
; Sequence 617, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-617

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCCTCT 3

Search completed: May 21, 2003, 08:44:51
Job time : 65 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:26:13 ; Search time 1092.5 Seconds
(without alignments)
88.945 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccucu 6

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gp_estl1:*
10: gp_est2:*
11: gp_hic:*
12: gp_est3:*
13: gp_est4:*
14: gp_est5:*
15: em_estfun:*
16: em_estom:*
17: gp_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 6 | 100.0 | 15 | 14 | B0511821 EST619236 |
| 2 | 6 | 100.0 | 16 | 14 | BQ789979 hage005a |
| 3 | 6 | 100.0 | 19 | 9 | A1569191 tr83f12.x |
| 4 | 6 | 100.0 | 19 | 17 | A2429998 1M0214F16 |
| 5 | 6 | 100.0 | 19 | 17 | A2475079 1M0293B17 |
| 6 | 6 | 100.0 | 19 | 17 | A2480905 1M0302N22 |

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 7 | 6 | 100.0 | 19 | 17 | A2509071 1M0351A21 |
| 8 | 6 | 100.0 | 19 | 17 | A2651803 1M0522N11 |
| 9 | 6 | 100.0 | 19 | 17 | A2659603 1M0537N06 |
| 10 | 6 | 100.0 | 19 | 17 | A2800646 2M0058B14 |
| 11 | 6 | 100.0 | 19 | 17 | A2834038 2M0116H01 |
| 12 | 6 | 100.0 | 20 | 14 | BQ789776 hage002a |
| 13 | 6 | 100.0 | 20 | 17 | A2309156 1M0013B09 |
| 14 | 6 | 100.0 | 20 | 17 | A2366535 1M0115M15 |
| 15 | 6 | 100.0 | 20 | 17 | A2615164 1M0444B24 |
| 16 | 6 | 100.0 | 20 | 17 | A2797468 2M0053P09 |
| 17 | 6 | 100.0 | 20 | 17 | A2807038 2M0069C06 |
| 18 | 6 | 100.0 | 21 | 17 | A2303734 1M0003O21 |
| 19 | 6 | 100.0 | 21 | 17 | A2317085 1M0035D07 |
| 20 | 6 | 100.0 | 21 | 17 | A2342282 1M0075H14 |
| 21 | 6 | 100.0 | 21 | 17 | A2650869 1M0521G14 |
| 22 | 6 | 100.0 | 22 | 9 | A1023917 ow69f05.s |
| 23 | 6 | 100.0 | 22 | 9 | A1128425 qc67f09.x |
| 24 | 6 | 100.0 | 22 | 9 | A1660937 wF20D06.x |
| 25 | 6 | 100.0 | 22 | 17 | A2787102 2M0033B03 |
| 26 | 6 | 100.0 | 22 | 17 | A2792724 2M0045E20 |
| 27 | 6 | 100.0 | 22 | 17 | A2803482 2M0063I13 |
| 28 | 6 | 100.0 | 22 | 17 | TA204A05P |
| 29 | 6 | 100.0 | 22 | 17 | TA294D03P |
| 30 | 6 | 100.0 | 23 | 17 | A2393604 1M0156C14 |
| 31 | 6 | 100.0 | 23 | 17 | A2779967 2M0016C22 |
| 32 | 6 | 100.0 | 23 | 17 | A2781980 2M0021N13 |
| 33 | 6 | 100.0 | 23 | 17 | A2783504 2M0025L17 |
| 34 | 6 | 100.0 | 23 | 17 | A2830526 2M0109M16 |
| 35 | 6 | 100.0 | 23 | 17 | A2954682 2M0220P23 |
| 36 | 6 | 100.0 | 23 | 17 | A2967993 2M0240J07 |
| 37 | 6 | 100.0 | 23 | 17 | BH812514 SALR 0618 |
| 38 | 6 | 100.0 | 24 | 17 | A2443047 1M0237I06 |
| 39 | 6 | 100.0 | 24 | 17 | A2463313 1M0272L03 |
| 40 | 6 | 100.0 | 24 | 17 | A2505865 1M0346C18 |
| 41 | 6 | 100.0 | 24 | 17 | A2780307 2M0017J04 |
| 42 | 6 | 100.0 | 24 | 17 | A2787390 2M0033J23 |
| 43 | 6 | 100.0 | 24 | 17 | A2789936 2M0038L17 |
| 44 | 6 | 100.0 | 24 | 17 | A2812679 2M0079H03 |
| 45 | 6 | 100.0 | 24 | 17 | A2936903 2M0193E20 |

ALIGNMENTS

RESULT 1
B0511821
LOCUS
DEFINITION
EST619236 Generation of a set of potato CDNA clones for microarray analyses mixed potato tissues Solanum tuberosum CDNA clone SYMH018 5' end, mRNA sequence.

ACCESSION
B0511821
VERSION
B0511821.1
KEYWORDS
SOURCE
ORGANISM

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanales; Solanales; Solanum.

REFERENCE
AUTHORS
Buell, C.R., Hartl, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karamycheva, S.A.
Generation of a set of potato CDNA clones for microarray analyses
Unpublished (2002)
Other ESTs: EST619237
Contact: Robin Buell

TITLE
JOURNAL
COMMENT
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@igf.org
This clone is available through the Research Genetics, contact the Research Genetics for further information 1-800-711-6195 or cdnaresgen.com
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1. 15
/organism="Solanum tuberosum"
/cultivar="Kennebec or Bintje"
/db_xref="taxon:4113"
/clone="STM018"
/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Combination of untreated and phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes , tubers, or roots."

BASE COUNT
ORIGIN
3 a 5 c 2 g 5 t

Query Match 100.0%; Score 6; DB 14; Length 15;
Best Local Similarity 66.7%; Pred. No. 7.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
Db 9 GCCTCT 14

RESULT 2
LOCUS B0789979 16 bp mRNA linear EST 30-JUL-2002
DEFINITION hage005a12 Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion annosum cDNA clone hage005a12, mRNA sequence.
ACCESSION B0789979
VERSION B0789979.1 GI:22004941
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 16)
Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
Selected expressed sequence tags of cDNA clones from the interaction of the root rot fungus (Heterobasidion annosum) with seedling roots of Scots pine (Pinus sylvestris)
Unpublished (2001)
CONTACT: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden
Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer
Location/Qualifiers
1. 16
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone="hage005a12"
/clone_lib="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
/note="Vector: pT-ADV; Site_1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of Heterobasidion annosum (FP5)."

BASE COUNT
ORIGIN
5 a 3 c 3 g 5 t

Query Match 100.0%; Score 6; DB 14; Length 16;
Best Local Similarity 66.7%; Pred. No. 8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
Db 7 GCCTCT 12

RESULT 3
LOCUS A1569191 19 bp mRNA linear EST 14-MAY-1999
DEFINITION t83f12.x1 NCI CGAP Pan1 Homo sapiens cDNA clone IMAGE:224943 3' similar to SW:PCEC.HUMAN P28074 PROTEASOME EPSILON CHAIN PRECURSOR ;, mRNA sequence.
ACCESSION A1569191
VERSION A1569191.1 GI:4532565
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 19)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps.rem@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bdrip/image/image.html

Trace considered overall poor quality
Insert length: 1117 Std Error: 0.00
Seq primer: -40UP from GIBCO
High quality sequence stop: 1
POLYA-NO.

FEATURES
source
Location/Qualifiers
1. 19
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:224943"
/clone_lib="NCI-CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: 01lgo dt. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT
ORIGIN
2 a 7 c 5 g 5 t

Query Match 100.0%; Score 6; DB 9; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
Db 8 GCCTCT 13

RESULT 4
LOCUS A2429998 19 bp DNA linear GSS 03-OCT-2000
DEFINITION M0214F16F Mouse 10kb plasmid UNGC1M library Mus musculus genomic clone UNGC1M0214F16 F, DNA sequence.
ACCESSION A2429998
VERSION A2429998.1 GI:10554011
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0214 row: F column: 16
Seq primer: CGTTGTAAACGACGCCACT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0214E16"
/clone_1lb="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[9b]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
5 a 3 c 9 g 2 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:
Db 11 GCCTCT 6

RESULT 5
A2475079 19 bp DNA linear GSS 04-OCT-2000
LOCUS A2475079/c
DEFINITION IM0293B17F Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0293B17 F, DNA sequence.
ACCESSION A2475079
VERSION A2475079.1 GI:106333204
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0293 row: B column: 17
Seq primer: CGTTGTAAACGACGCCACT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0293B17"
/clone_1lb="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[9b]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 4 c 9 g 0 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:
Db 18 GCCTCT 13

RESULT 6
A2480905 19 bp DNA linear GSS 04-OCT-2000
LOCUS A2480905
DEFINITION IM0302N22R Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0302N22 R, DNA sequence.
ACCESSION A2480905
VERSION A2480905.1 GI:10642066
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112 USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 22
Seq primer: CACACAGAAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0302N2"
/clone_1id="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD2 (g1147321149b) (AF129072.1) a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
1 a 5 c 9 g 4 t

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|
Db 7 GCCTCT 12

RESULT 7
A2309071/c 19 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0351A21R Mouse 10kb plasmid UGC1M library Mus musculus genomic.
DEFINITION clone UGC1M0351A21 R, DNA sequence.
ACCESSION A2309071
VERSION A2309071.1 GI:10690387
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112 USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0351 row: A column: 21
Seq primer: CACACAGAAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0351A21"
/clone_1id="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD2 (g1147321149b) (AF129072.1) a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
4 a 4 c 9 g 2 t

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|
Db 13 GCCTCT 8

RESULT 8
A2651803 19 bp DNA linear GSS 14-DEC-2000
LOCUS A2651803
DEFINITION 1M0522N11R Mouse 10kb plasmid UGC1M library Mus musculus genomic.
DEFINITION clone UGC1M0522N11 R, DNA sequence.
ACCESSION A2651803
VERSION A2651803.1 GI:11787672
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0522 row: N column: 11
Seq primer: CACACAGGAACGACGATGACCC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0522N11"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 7 c 5 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:|

Db 7 GCCTCT 12

RESULT 9
AZ659603/c 19 bp DNA linear GSS 14-DEC-2000
LOCUS
DEFINITION 1M0537N06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0537N06 F, DNA sequence.
ACCESSION AZ659603
VERSION AZ659603.1 GI:11796749
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0537 row: N column: 06
Seq primer: CGTTGTAAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0537N06"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
7 a 2 c 7 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:|

Db 12 GCCTCT 7

RESULT 10
AZ800646/c 19 bp DNA linear GSS 16-FEB-2001
LOCUS
DEFINITION 2M0058H14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0058H14 R, DNA sequence.
ACCESSION AZ800646
VERSION AZ800646.1 GI:12952969
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0058 row: H column: 14
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0058H14"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114[gb]AF129072.1), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 4 c 7 g 5 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:
Db 7 GCCTCT 2

RESULT 11
AZ834038 19 bp DNA linear GSS 20-FEB-2001
LOCUS 2M016H01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M016H01 R, DNA sequence.
ACCESSION AZ834038
VERSION AZ834038.1 GI:13003946
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0116 row: H column: 01
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0116H01"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114[gb]AF129072.1), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 2 c 7 g 4 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 8.6e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:
Db 6 GCCTCT 1

RESULT 12
BQ789776 20 bp mRNA linear EST 30-JUL-2002
LOCUS BQ789776
DEFINITION hage002a10 Heterobasidion annosum - Scots pine infection stage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage002a10, mRNA sequence.
ACCESSION BQ789776
VERSION BQ789776.1 GI:22004738
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 20)

AUTHORS Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
TITLE Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (*Heterobasidion annosum*) with seedling roots of Scots pine (*Pinus sylvestris*)

JOURNAL Unpublished (2001)
COMMENT Contact: Fred O. Asiegbu
 Dept. of Forest Mycology & Pathology
 Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden
 Tel: +46 18 67 15 98
 Fax: +46 18 30 92 45
 Email: Fred.Asiegbu@mykopat.slu.se
 Seq primer: T7 primer.

FEATURES
 source Location/Qualifiers

1..20
 /organism="Pinus sylvestris/Heterobasidion annosum"
 /db_xref="taxon:169015"
 /clone_id="hage002a10"
 /clone_1lb="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
 /dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
 /note="Vector: pT-Adv; Site_1: EcoRI. The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of *Heterobasidion annosum* (FP5)."

BASE COUNT 4 a 5 c 3 g 7 t 1 others
ORIGIN

Query Match 100.0%; Score 6; DB 14; Length 20;
Best Local Similarity 66.7%; Pred. No. 8.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 |||:|:
Db 12 GCGTCT 17

RESULT 13
AZ309156 20 bp DNA linear GSS 29-SEP-2000
LOCUS IM0013B09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0013B09 F, DNA sequence.
ACCESSION AZ309156
VERSION AZ309156.1 GI:10349862
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Baecorn, T., Duval, B., Hamll, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0013 row: B column: 09
 Seq primer: CGTCTAAACGACGCCACG
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers

FEATURES
 source 1..20

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0013B09"
 /clone_1lb="Mouse 10kb plasmid UUGC1M library"
 /sex="male"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1) a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 11 c 4 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 66.7%; Pred. No. 8.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 |||:|:
Db 3 GCGTCT 8

RESULT 14
AZ366535 20 bp DNA linear GSS 02-OCT-2000
LOCUS IM0115M15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0115M15 R, DNA sequence.
ACCESSION AZ366535
VERSION AZ366535.1 GI:10480235
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Baecorn, T., Duval, B., Hamll, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0115 row: M column: 15
 Seq primer: CACACAGCAACACGATGACC
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers

FEATURES
 source 1..20

```

/organism="Mus musculus"
/strain="C57BL/6j"
/db_xref="taxon:10090"
/clone="UUC1M0115M15"
/clone.lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6j (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

BASE COUNT
ORIGIN

5 a 6 c 3 g 6 t

Query Match

Best Local Similarity 66.7%; Score 6; DB 17; Length 20;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:

DB 14 GCCTCT 19

RESULT 15
A2615164

LOCUS 20 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0444B24F Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M044B24 F, DNA sequence.

ACCESSION A2615164

VERSION A2615164.1 GI:11737354

KEYWORDS GSS.

SOURCE

ORGANISM

house mouse.

Mus musculus

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0444 row: B column: 24
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1. .20

TITLE

JOURNAL
COMMENT

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

FEATURES
source

```

/organism="Mus musculus"
/strain="C57BL/6j"
/db_xref="taxon:10090"
/clone="UUC1M044B24"
/clone.lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6j (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

BASE COUNT
ORIGIN

3 a 7 c 5 g 5 t

Query Match

Best Local Similarity 66.7%; Score 6; DB 17; Length 20;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:

DB 7 GCCTCT 12

Search completed: May 21, 2003, 08:37:45
Job time : 1092.5 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:56:27 : Search time 333.5 Seconds

(without alignments)
523.589 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6

Sequence: 1 gccucc 6

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pin:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| C 1 | 6 | 100.0 | 10 | 6 | AR128987 | AR128987 Sequence |
| C 2 | 6 | 100.0 | 10 | 6 | AR128988 | AR128988 Sequence |
| C 3 | 6 | 100.0 | 10 | 6 | AR128989 | AR128989 Sequence |
| C 4 | 6 | 100.0 | 10 | 6 | AR128990 | AR128990 Sequence |
| C 5 | 6 | 100.0 | 10 | 6 | AR128991 | AR128991 Sequence |
| C 6 | 6 | 100.0 | 10 | 6 | AX152349 | AX152349 Sequence |
| C 7 | 6 | 100.0 | 10 | 6 | AX152410 | AX152410 Sequence |
| C 8 | 6 | 100.0 | 10 | 6 | AX152849 | AX152849 Sequence |
| C 9 | 6 | 100.0 | 10 | 6 | AX152886 | AX152886 Sequence |
| C 10 | 6 | 100.0 | 10 | 6 | AX153384 | AX153384 Sequence |
| C 11 | 6 | 100.0 | 10 | 6 | AX153513 | AX153513 Sequence |
| C 12 | 6 | 100.0 | 10 | 6 | E16890 | E16890 DNA sequence |
| C 13 | 6 | 100.0 | 11 | 6 | A91501 | A91501 Sequence 28 |
| C 14 | 6 | 100.0 | 11 | 6 | A91503 | A91503 Sequence 30 |
| C 15 | 6 | 100.0 | 11 | 6 | AX470495 | AX470495 Sequence |
| C 16 | 6 | 100.0 | 11 | 6 | AX470740 | AX470740 Sequence |
| C 17 | 6 | 100.0 | 11 | 6 | AX470747 | AX470747 Sequence |
| C 18 | 6 | 100.0 | 11 | 6 | AX471502 | AX471502 Sequence |
| C 19 | 6 | 100.0 | 11 | 6 | AX471630 | AX471630 Sequence |
| C 20 | 6 | 100.0 | 11 | 6 | AX471677 | AX471677 Sequence |
| C 21 | 6 | 100.0 | 11 | 6 | AX471805 | AX471805 Sequence |
| C 22 | 6 | 100.0 | 12 | 6 | A06058 | A06058 Synthetic P |
| C 23 | 6 | 100.0 | 12 | 6 | A06059 | A06059 Synthetic P |
| C 24 | 6 | 100.0 | 12 | 6 | A15123 | A15123 Nucleotide |
| C 25 | 6 | 100.0 | 12 | 6 | A16601 | A16601 Nucleotide |
| C 26 | 6 | 100.0 | 12 | 6 | A16602 | A16602 Nucleotide |
| C 27 | 6 | 100.0 | 12 | 6 | A47656 | A47656 Sequence 16 |
| C 28 | 6 | 100.0 | 12 | 6 | A61520 | A61520 Sequence 89 |
| C 29 | 6 | 100.0 | 12 | 6 | A91489 | A91489 Sequence 16 |
| C 30 | 6 | 100.0 | 12 | 6 | AR027874 | AR027874 Sequence |
| C 31 | 6 | 100.0 | 12 | 6 | AR075465 | AR075465 Sequence |
| C 32 | 6 | 100.0 | 12 | 6 | AR075468 | AR075468 Sequence |
| C 33 | 6 | 100.0 | 12 | 6 | AR101000 | AR101000 Sequence |
| C 34 | 6 | 100.0 | 12 | 6 | AR137925 | AR137925 Sequence |
| C 35 | 6 | 100.0 | 12 | 6 | AR153916 | AR153916 Sequence |
| C 36 | 6 | 100.0 | 12 | 6 | AR153919 | AR153919 Sequence |
| C 37 | 6 | 100.0 | 12 | 6 | AR167701 | AR167701 Sequence |
| C 38 | 6 | 100.0 | 12 | 6 | AR178533 | AR178533 Sequence |
| C 39 | 6 | 100.0 | 12 | 6 | AR178536 | AR178536 Sequence |
| C 40 | 6 | 100.0 | 12 | 6 | AR199100 | AR199100 Sequence |
| C 41 | 6 | 100.0 | 12 | 6 | AX233626 | AX233626 Sequence |
| C 42 | 6 | 100.0 | 12 | 6 | E29585 | E29585 Method for |
| C 43 | 6 | 100.0 | 12 | 6 | E38691 | E38691 Method and |
| C 44 | 6 | 100.0 | 12 | 6 | E64117 | E64117 Method for |
| C 45 | 6 | 100.0 | 12 | 6 | I07918 | I07918 Sequence 30 |

ALIGNMENTS

RESULT 1

AR128987/c

LOCUS AR128987 10 bp DNA Linear PAT 16-MAY-2001

DEFINITION Sequence 2 from patent US 6183966.

ACCESSION AR128987

VERSION AR128987.1 GI:14116649

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 10)

AUTHORS Gray,D.M. and Clark,C.L.

TITLE Apparatus and method for selectively ranking sequences for antisense targeting

JOURNAL Patent: US 6183966-A 2 06-FEB-2001;

FEATURES location/Qualifiers
 source 1..10
 BASE COUNT 2 a 3 c 4 g 1 t
 ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 Db 6 GCCTCT 1

RESULT 2
 LOCUS ARI28988 10 bp DNA PAT 16-MAY-2001
 DEFINITION Sequence 3 from patent US 6183966.
 ACCESSION ARI28988
 VERSION ARI28988.1 GI:14116650
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Gray,D.M. and Clark,C.L.
 TITLE Apparatus and method for selectively ranking sequences for
 antisense targeting
 JOURNAL Patent: US 6183966-A 3 06-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..10

BASE COUNT 2 a 3 c 5 g 0 t
 ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 Db 7 GCCTCT 2

RESULT 3
 LOCUS ARI28989 10 bp DNA PAT 16-MAY-2001
 DEFINITION Sequence 4 from patent US 6183966.
 ACCESSION ARI28989
 VERSION ARI28989.1 GI:14116651
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Gray,D.M. and Clark,C.L.
 TITLE Apparatus and method for selectively ranking sequences for
 antisense targeting
 JOURNAL Patent: US 6183966-A 4 06-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..10

BASE COUNT 3 a 3 c 4 g 0 t
 ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 Db 8 GCCTCT 3

RESULT 4
 LOCUS ARI28990 10 bp DNA PAT 16-MAY-2001
 DEFINITION Sequence 5 from patent US 6183966.
 ACCESSION ARI28990
 VERSION ARI28990.1 GI:14116652
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Gray,D.M. and Clark,C.L.
 TITLE Apparatus and method for selectively ranking sequences for
 antisense targeting
 JOURNAL Patent: US 6183966-A 5 06-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..10

BASE COUNT 4 a 2 c 4 g 0 t
 ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 Db 9 GCCTCT 4

RESULT 5
 LOCUS ARI28991 10 bp DNA PAT 16-MAY-2001
 DEFINITION Sequence 6 from patent US 6183966.
 ACCESSION ARI28991
 VERSION ARI28991.1 GI:14116653
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Gray,D.M. and Clark,C.L.
 TITLE Apparatus and method for selectively ranking sequences for
 antisense targeting
 JOURNAL Patent: US 6183966-A 6 06-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..10

BASE COUNT 4 a 2 c 4 g 0 t
 ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
 Db 10 GCCTCT 5

RESULT 6
 LOCUS AX152349 10 bp DNA PAT 22-JUN-2001
 DEFINITION Sequence 264 from patent W00138577.
 ACCESSION AX152349
 VERSION AX152349.1 GI:14534000
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 264 31-MAY-2001;
FEATURES Location/Qualifiers
    source
        1..10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 0 a 5 c 2 g 3 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 1 GCCTCT 6

RESULT 7
AX152410 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 325 from Patent W00138577.
DEFINITION AX152410
ACCESSION AX152410.1 GI:14534061
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 325 31-MAY-2001;
FEATURES Location/Qualifiers
    source
        1..10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 1 a 5 c 2 g 2 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 1 GCCTCT 6

RESULT 8
AX152849 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 764 from Patent W00138577.
DEFINITION AX152849
ACCESSION AX152849
VERSION AX152849.1 GI:14534500
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 764 31-MAY-2001;
FEATURES Location/Qualifiers
    source
        1..10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 0 a 5 c 2 g 3 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 1 GCCTCT 6

RESULT 9
AX152886 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 801 from Patent W00138577.
DEFINITION AX152886
ACCESSION AX152886
VERSION AX152886.1 GI:14534537
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 801 31-MAY-2001;
FEATURES Location/Qualifiers
    source
        1..10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 5 a 1 c 3 g 1 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 8 GCCTCT 3

RESULT 10
AX153384 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 1299 from Patent W00138577.
DEFINITION AX153384
ACCESSION AX153384
VERSION AX153384.1 GI:14535035
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 1299 31-MAY-2001;
FEATURES Location/Qualifiers
    source
        1..10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN

Query Match
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 1 GCCTCT 6
```

Db 2 GCCTCT 7

RESULT 11
AX153513
LOCUS AX153513 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1428 from Patent WO0138577.
ACCESSION AX153513
VERSION AX153513.1 GI:14535164
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Velculescu, V.E., Vogelstein, B. and Kinzler, K.W.
TITLE Human transcriptomes
JOURNML Patent: WO 0138577-A 1428 31-MAY-2001;
The Johns Hopkins University (US)
Location/Qualifiers
FEATURES
1. .10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 0 a 4 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
111:1
Db 1 GCCTCT 6

RESULT 12
E16890/c
LOCUS E16890 10 bp DNA linear PAT 28-JUL-1999
DEFINITION DNA sequence required for efficient protein transcription in
Brevibacterium flavum.
ACCESSION E16890
VERSION E16890.1 GI:5711573
KEYWORDS JP 1998229881-A/31.
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 (bases 1 to 10)
AUTHORS Kobayashi, M., Man, T. and Yugawa, H.
TITLE DNA HAVING SEQUENCE CAPABLE OF EFFICIENTLY TRANSLATING PROTEIN IN
JOURNML CORYNEFORM BACTERIA
PATENT: JP 1998229881-A 31 02-SEP-1998;
MITSUBISHI CHEM CORP
OS Brevibacterium flavum
PN JP 1998229881-A/31
PD 02-SEP-1998
PF 19-FEB-1997 JP 1997035338
PI KOBAYASHI MIKI, MAN TOMOKO, YUGAWA HIDEAKI
PC C12N15/09.C07H21/04.C12N1/21//C12N9/38.C12Q1/68.(C12N15/09, PC
C12N1/19)
PC (C12N1/21,C12N1/33),(C12N9/38,C12N1/19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1. .10
/strain="MF-233",
/organism="Brevibacterium flavum" FT
/location/Qualifiers
1. .10
/organism="Corynebacterium glutamicum"

FEATURES
source

BASE COUNT 3 a 1 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
111:1
Db 9 GCCTCT 4

RESULT 13
A91501
LOCUS A91501 11 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 28 from Patent WO9824928.
ACCESSION A91501
VERSION A91501.1 GI:6740456
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 11)
AUTHORS Pallisgaard, N. and Hokland, P.
TITLE DETECTION OF CHROMOSOMAL ABNORMALITIES
JOURNML Patent: WO 9824928-A 28 11-JUN-1998;
PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
FEATURES
1. .11
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1 a 5 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
111:1
Db 5 GCCTCT 10

RESULT 14
A91503
LOCUS A91503 11 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 30 from Patent WO9824928.
ACCESSION A91503
VERSION A91503.1 GI:6740458
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 11)
AUTHORS Pallisgaard, N. and Hokland, P.
TITLE DETECTION OF CHROMOSOMAL ABNORMALITIES
JOURNML Patent: WO 9824928-A 30 11-JUN-1998;
PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
FEATURES
1. .11
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 0 a 5 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
111:1
Db 5 GCCTCT 10

RESULT 15
AX470495/c AX470495 11 bp DNA linear PAT 09-AUG-2002
LOCUS Sequence 72 from Patent WO02053773.
DEFINITION AX470495
ACCESSION AX470495
VERSION AX470495.1 GI:22205620
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Hofmann, K., Conradt, M. and Petersohn, D.
TITLE Method for determining skin stress or skin ageing in vitro
JOURNAL Patent: WO 02053773-A 72 11-JUL-2002;
HENKEL KGAA (DE)
FEATURES
source location/Qualifiers
1..11
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 6 a 2 c 3 g 0 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 66.7%; Pred. No. 2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
|||:|
Db 8 GCCTCT 3
Search completed: May 21, 2003, 05:12:46
Job time : 334 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:52:32 ; Search time 156 Seconds

(without alignments)
86.615 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccucc 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_101002:*

| | |
|-----|--|
| 1: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:* |
| 2: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:* |
| 3: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:* |
| 4: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:* |
| 5: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:* |
| 6: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:* |
| 7: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:* |
| 8: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:* |
| 9: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:* |
| 10: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:* |
| 11: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:* |
| 12: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:* |
| 13: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:* |
| 14: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:* |
| 15: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:* |
| 16: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:* |
| 17: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:* |
| 18: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:* |
| 19: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:* |
| 20: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:* |
| 21: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:* |
| 22: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:* |
| 23: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:* |
| 24: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 6 | 100.0 | 8 | 20 | AAAX29578 |
| 2 | 6 | 100.0 | 8 | 21 | AAAB0736 |
| 3 | 6 | 100.0 | 9 | 24 | ABQ71469 |
| 4 | 6 | 100.0 | 9 | 24 | ABQ71786 |
| 5 | 6 | 100.0 | 9 | 24 | ABQ71788 |
| 6 | 6 | 100.0 | 9 | 24 | ABQ71921 |
| 7 | 6 | 100.0 | 9 | 24 | ABQ71982 |
| 8 | 6 | 100.0 | 9 | 24 | ABQ72173 |
| 9 | 6 | 100.0 | 9 | 24 | ABQ72175 |

| | | | | | | | |
|---|----|---|-------|----|----|----------|--------------------|
| C | 10 | 6 | 100.0 | 10 | 16 | AAQ96863 | HIV-1 NL4-3 nef ge |
| C | 11 | 6 | 100.0 | 10 | 16 | AAQ96864 | HIV-1 NL4-3 nef ge |
| C | 12 | 6 | 100.0 | 10 | 16 | AAQ96865 | HIV-1 NL4-3 nef ge |
| C | 13 | 6 | 100.0 | 10 | 16 | AAQ96866 | HIV-1 NL4-3 nef ge |
| C | 14 | 6 | 100.0 | 10 | 16 | AAQ96867 | HIV-1 NL4-3 nef ge |
| C | 15 | 6 | 100.0 | 10 | 20 | AAQ86201 | SAGE tag used to 1 |
| C | 16 | 6 | 100.0 | 10 | 20 | AAQ00518 | Antisense oligonuc |
| C | 17 | 6 | 100.0 | 10 | 20 | AAQ00519 | Antisense oligonuc |
| C | 18 | 6 | 100.0 | 10 | 20 | AAQ00522 | Antisense oligonuc |
| C | 19 | 6 | 100.0 | 10 | 20 | AAQ00520 | Antisense oligonuc |
| C | 20 | 6 | 100.0 | 10 | 20 | AAQ00521 | Antisense oligonuc |
| C | 21 | 6 | 100.0 | 10 | 21 | AAQ56532 | Human macrophage g |
| C | 22 | 6 | 100.0 | 10 | 21 | AAQ77725 | Human dendritic ce |
| C | 23 | 6 | 100.0 | 10 | 21 | AAQ77746 | Human dendritic ce |
| C | 24 | 6 | 100.0 | 10 | 21 | AAQ77927 | Human dendritic ce |
| C | 25 | 6 | 100.0 | 10 | 21 | AAQ78138 | Human dendritic ce |
| C | 26 | 6 | 100.0 | 10 | 21 | AAQ78540 | Human dendritic ce |
| C | 27 | 6 | 100.0 | 10 | 21 | AAQ78549 | Human dendritic ce |
| C | 28 | 6 | 100.0 | 10 | 21 | AAQ78652 | Human dendritic ce |
| C | 29 | 6 | 100.0 | 10 | 21 | AAQ78858 | Human dendritic ce |
| C | 30 | 6 | 100.0 | 10 | 21 | AAQ79181 | Human dendritic ce |
| C | 31 | 6 | 100.0 | 10 | 21 | AAQ79724 | Human colon tumour |
| C | 32 | 6 | 100.0 | 10 | 21 | AAQ79825 | Human lung prefere |
| C | 33 | 6 | 100.0 | 10 | 21 | AAQ79884 | Human dendritic ce |
| C | 34 | 6 | 100.0 | 10 | 21 | AAQ81089 | Metastatic breast |
| C | 35 | 6 | 100.0 | 10 | 21 | AAQ81276 | Metastatic breast |
| C | 36 | 6 | 100.0 | 10 | 21 | AAQ81485 | Metastatic breast |
| C | 37 | 6 | 100.0 | 10 | 21 | AAQ81751 | Metastatic breast |
| C | 38 | 6 | 100.0 | 10 | 21 | AAQ81798 | Metastatic breast |
| C | 39 | 6 | 100.0 | 10 | 21 | AAQ82217 | Metastatic breast |
| C | 40 | 6 | 100.0 | 10 | 21 | AAQ82317 | Metastatic breast |
| C | 41 | 6 | 100.0 | 10 | 21 | AAQ82473 | Metastatic breast |
| C | 42 | 6 | 100.0 | 10 | 21 | AAQ82783 | Metastatic breast |
| C | 43 | 6 | 100.0 | 10 | 21 | AAQ82938 | Metastatic breast |
| C | 44 | 6 | 100.0 | 10 | 21 | AAQ83462 | Metastatic breast |
| C | 45 | 6 | 100.0 | 10 | 21 | AAQ84090 | Metastatic breast |

ALIGNMENTS

RESULT 1

| | | |
|----|---|----------------------|
| ID | AAAX29578 | standard; DNA: 8 BP. |
| AC | AAAX29578: | |
| XX | | |
| DT | 03-JUN-1999 | (first entry) |
| XX | | |
| DE | Primer for human G-protein coupled receptor genes. | |
| XX | | |
| KW | Nucleic acid amplification; nuclear receptor; G-protein coupled receptor; | |
| KW | apoptosis; DNA repair; DNA replication; plant biology; agriculture; | |
| KW | human; veterinary medicine; reproduction; microbiology; hybridisation; | |
| KW | environmental science; DNA fingerprinting; PCR primer; ss. | |
| XX | | |
| OS | Synthetic. | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO9911823-A2. | |
| XX | | |
| PD | 11-MAR-1999. | |
| XX | | |
| PF | 04-SEP-1998; | 98WO-US18392. |
| XX | | |
| PR | 05-SEP-1997; | 97US-0925816. |
| XX | | |
| PA | (KIMM-) KIMMEL CANCER CENT SIDNEY. | |
| XX | | |
| PI | McClelland M, Pesole G; | |
| XX | | |
| DR | WPI; 1999-205200/17. | |
| XX | | |

PT Subset of primers able to amplify group of related sequences
 XX
 PS Claim 21; Page 75; 92pp; English.
 XX
 CC The invention provides primers (AAx29501-X29679) for identifying
 CC sequences encoding structurally or functionally related proteins such as
 CC nuclear or G-protein coupled receptors, apoptosis-related or DNA
 CC repair/replication proteins. The identified sequences are broadly useful
 CC in plant biology, agriculture, human or veterinary medicine,
 CC reproduction, microbiology or environmental science, e.g. to study
 CC expression of nuclear receptors at different stages of tissue development
 CC or after treatment with particular drugs. It is also used for DNA
 CC fingerprinting (to generate products useful for differential
 CC hybridisation), or, where a 3'-anchor primer is used, to isolate the
 CC 3'-ends of mRNA sequences. Sequences AAx29526-X29583 represent claimed
 CC primers specific for human G-protein coupled receptor genes.
 XX
 SQ Sequence 8 BP; 0 A; 4 C; 2 G; 2 T; 0 other;
 Query Match 100.0%; Score 6; DB 20; Length 8;
 Best Local Similarity 66.7%; Pred. No. 2.7e+08;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCUCU 6
 1111:
 Db 1 GCCTCT 6
 1111:
 RESULT 2
 ID AAA80736 standard; DNA; 8 BP.
 AC AAA80736;
 XX
 XX
 DE 24-NOV-2000 (first entry)
 DT
 XX
 DE A. thaliana primer walking octamer SEQ ID NO: 49.
 KW
 XX
 KW Primer walking; octamer; primer; DNA sequencing; PCR; ss.
 OS
 XX
 OS Arabidopsis thaliana.
 XX
 PN US6083695-A.
 XX
 PD 04-JUL-2000.
 PD
 XX
 PF 21-MAY-1997; 97US-0859954.
 PF
 XX
 PR 15-APR-1996; 96US-0632782.
 PR
 XX
 PA (UYHO-) UNIV HOUSTON.
 PA
 PA (HARD/) HARDIN S H.
 PI
 PI Hardin PE, Hardin SH, Homeyouni R;
 DR
 DR WPI: 2000-474852/41.
 PT
 PT Sequencing an unknown DNA molecule for the polymerase chain reaction
 PT and other primer processes comprises primer walking of octamer
 PT oligonucleotides
 PS
 PS Example 8; Column 49-50; 161pp; English.
 XX
 CC This invention describes a novel method for sequencing an unknown DNA
 CC molecule which comprises selecting a library primer from an octamer
 CC oligonucleotide library consisting of 48 8-bp sequences and
 CC corresponding complementary sequences, where the library primer is
 CC complementary to a known sequence adjacent to the unknown sequence or
 CC is complementary to a sequence in a known extension product. The method
 CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
 CC which make use of primers. The octamers are used to identify coding
 CC sequences. Primer walking using the octamer libraries is advantageous
 CC over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
 CC directed and methodical approach. AAA80688-AB1253 represent the octamer
 CC primers used in the primer walking method of the invention.
 XX
 SQ Sequence 8 BP; 0 A; 3 C; 1 G; 4 T; 0 other;
 Query Match 100.0%; Score 6; DB 21; Length 8;
 Best Local Similarity 66.7%; Pred. No. 2.7e+08;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCUCU 6
 1111:
 Db 3 GCCTCT 8
 1111:
 RESULT 3
 ID ABQ71469/C
 AC ABQ71469 standard; DNA; 9 BP.
 AC ABQ71469;
 XX
 XX
 DE 28-AUG-2002 (first entry)
 DT
 XX
 DE Zinc finger protein related oligonucleotide target SEQ ID NO:588.
 DE
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
 KW
 XX
 OS Homo sapiens.
 OS
 OS Synthetic.
 XX
 PN WO200242459-A2.
 PN
 PD 30-MAY-2002.
 PD
 XX
 PF 20-NOV-2001; 2001WO-US43438.
 PF
 XX
 PR 20-NOV-2000; 2000US-0716637.
 PR
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 PA
 PI Liu Q;
 PI
 PI WPI: 2002-500284/53.
 DR
 DR
 PT
 PT New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering.
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus
 PS
 PS Example 1; Page 44; 81pp; English.
 XX
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 CC that it binds to the S2 target subsite, and selecting the F3 zinc
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.


```
XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;
SQ Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.4e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 8 GCCTCT 3

RESULT 4
ABQ71786/c
ID ABQ71786 standard; DNA; 9 BP.
XX
XX ABQ71786;
AC
XX
XX 28-AUG-2002 (first entry)
DT
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2084.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX
XX Homo sapiens.
OS
XX Synthetic.
OS
XX WO200242459-A2.
PN
XX
XX 30-MAY-2002.
PD
XX
XX 20-NOV-2001; 2001MO-US43438.
PF
XX
XX 20-NOV-2000; 2000US-0716637.
PR
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA
XX
XX Liu Q;
PI
XX
XX WPI; 2002-500284/53.
DR
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
XX
XX Example 1; Page 55; 81pp; English.
PS
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (i) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;
```

```
XX Query Match 100.0%; Score 6; DB 24; Length 9;
SQ Best Local Similarity 66.7%; Pred. No. 2.4e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
Db 8 GCCTCT 3

RESULT 5
ABQ71788/c
ID ABQ71788 standard; DNA; 9 BP.
XX
XX ABQ71788;
AC
XX
XX 28-AUG-2002 (first entry)
DT
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2086.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX
XX Homo sapiens.
OS
XX Synthetic.
OS
XX WO200242459-A2.
PN
XX
XX 30-MAY-2002.
PD
XX
XX 20-NOV-2001; 2001MO-US43438.
PF
XX
XX 20-NOV-2000; 2000US-0716637.
PR
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA
XX
XX Liu Q;
PI
XX
XX WPI; 2002-500284/53.
DR
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
XX
XX Example 1; Page 55; 81pp; English.
PS
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (i) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;
```

OY 1 GCCUCU 6
||||:
DB 8 GCCTCT 3

RESULT 6

ABQ71921/c
ID ABQ71921 standard; DNA: 9 BP.

AC ABQ71921;

DE 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2219.

DE Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN WO200242459-A2.

PD 30-MAY-2002.

PF 20-NOV-2001; 2001WO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus

PS Example 1; Page 58; 81pp; English.

CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3',-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. NO. 2.4e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:
DB 8 GCCTCT 3

DB 8 GCCTCT 3

RESULT 7

ABQ71982/c
ID ABQ71982 standard; DNA: 9 BP.

AC ABQ71982;

DE 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2280.

DE Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN WO200242459-A2.

PD 30-MAY-2002.

PF 20-NOV-2001; 2001WO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus

PS Example 1; Page 59; 81pp; English.

CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3',-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. NO. 2.4e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
||||:
DB 8 GCCTCT 3

```
RESULT 8
AB072173/c
ID AB072173 standard; DNA; 9 BP.
XX
XX
AC AB072173;
XX
XX
DT 28-AUG-2002 (first entry)
XX
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2471.
XX
XX
KM Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX
OS Homo sapiens.
XX
XX
SS Synthetic.
XX
XX
PN WO200242459-A2.
XX
XX
PD 30-MAY-2002.
XX
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
XX
PR 20-NOV-2000; 2000US-0716637.
XX
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX
PI Liu Q;
XX
XX
DR WPI; 2002-500284/53.
XX
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus.
XX
XX
PS Example 1; Page 63; 81pp; English.
XX
XX
CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.4e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3
XX
RESULT 9
AB072175/c
ID AB072175 standard; DNA; 9 BP.
XX
```

```
XX
XX
AC AB072175;
XX
XX
DT 28-AUG-2002 (first entry)
XX
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2473.
XX
XX
KM Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX
OS Homo sapiens.
XX
XX
SS Synthetic.
XX
XX
PN WO200242459-A2.
XX
XX
PD 30-MAY-2002.
XX
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
XX
PR 20-NOV-2000; 2000US-0716637.
XX
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX
PI Liu Q;
XX
XX
DR WPI; 2002-500284/53.
XX
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus.
XX
XX
PS Example 1; Page 63; 81pp; English.
XX
XX
CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.4e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3
XX
RESULT 10
AA096863/c
ID AA096863 standard; DNA; 10 BP.
XX
XX
AC AA096863;
XX
XX
```

```

DT 26-MAR-1996 (first entry)
XX
XX HIV-1 NL4-3 nef gene nucleotide deletion 458.
DE
XX
XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
KM
XX
OS Human immunodeficiency virus type 1.
XX
XX W09521912-A1.
PN
XX 17-AUG-1995.
PD
XX
XX 14-FEB-1995; 95WO-AU00063.
PF
XX
XX 23-DEC-1994; 94AU-0000284.
PR 14-FEB-1994; 94AU-0003864.
PR 21-FEB-1994; 94AU-0004002.
XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI; 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
PS
XX
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AA096406-097018) from the nef gene and/or
CC 1 or more decanucleotides (AA097019-097166) from the LTR region; the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
CC
XX
XX Sequence 10 BP; 4 A; 1 C; 4 G; 1 T; 0 other;
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
Db 10 CCTCT 5

```

```

XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI; 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
PS
XX
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AA096406-097018) from the nef gene and/or
CC 1 or more decanucleotides (AA097019-097166) from the LTR region; the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
CC
XX
XX Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
Db 9 CCTCT 4

```

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 66.7%; Pred. No. 2.7e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
 |||:|:
 Db 8 GCCTCT 3

RESULT 13
 AA096866/c
 ID AA096866 standard; DNA; 10 BP.

XX
 AC AA096866;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 461.

XX
 KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX
 PD 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

XX
 PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

XX
 PR 21-FEB-1994; 94AU-0004002.

XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX
 PA (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX
 DR WPI; 1995-293115/38.

XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

XX
 PS Claim 13; Page 194; 301pp; English.

XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deannucleotides (AA096406-Q97018) from the nef gene and/or

CC 1 or more deannucleotides (AA097019-Q97166) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene

CC (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX

SQ Sequence 10 BP; 5 A; 2 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 66.7%; Pred. No. 2.7e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
 |||:|:
 Db 7 GCCTCT 2

RESULT 14

AA096867/c
 ID AA096867 standard; DNA; 10 BP.

XX
 AC AA096867;

XX
 DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 462.

XX
 KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX
 PD 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

XX
 PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

XX
 PR 21-FEB-1994; 94AU-0004002.

XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX
 PA (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX
 DR WPI; 1995-293115/38.

XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

XX
 PS Claim 13; Page 194; 301pp; English.

XX
 CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deannucleotides (AA096406-Q97018) from the nef gene and/or

CC 1 or more deannucleotides (AA097019-Q97166) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene

CC (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX

SQ Sequence 10 BP; 4 A; 2 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 66.7%; Pred. No. 2.7e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
 |||:|:
 Db 6 GCCTCT 1

RESULT 15
 AAX86201
 ID AAX86201 standard; DNA; 10 BP.

XX
 AC AAX86201;

XX
 DT 22-SEP-1999 (first entry)

XX
 DE SAGE tag used to identify transcripts which are enhanced by p53.

XX
 KM p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity;
 XX neoplastic; p53 binding site; PIG-3 promoter; SAGE tag; ss.
 OS Homo sapiens.
 XX
 PN WO9914356-A2.
 XX
 PD 25-MAR-1999.

```

XX 17-SEP-1998; 98WO-US19300.
PF
XX
PR 30-MAR-1998; 98US-0079817.
PR 17-SEP-1997; 97US-0059153.
XX
PA (UWJO ) UNIV JOHNS HOPKINS.
XX
PI Kinzler KW, Polyak K, Vogelstein B;
XX
DR WPL; 1999-443793/37.
XX
PT Use of p53 transcription tags to determine p53 status in, e.g.
PT cancer diagnosis
PS
XX Example 1; Page 25; 73pp; English.
XX
CC The specification describes the use of p53 transcription tags for
CC developing products to determine p53 status, to diagnose cancer
CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
CC A method for diagnosing cancer or determining p53 status in a sample
CC suspected for being neoplastic comprises comparing the level of
CC transcription of an RNA transcript in a first sample (s1) of a first
CC tissue (t1) to the level of transcription of the transcript in a second
CC sample (s2) of a second tissue (s2), where s1 is suspected of being
CC neoplastic and s2 is a normal human tissue (of the same type) and the
CC transcript is identified by a tag; and categorizing s1 as neoplastic
CC or as having a mutant p53 when transcription is found to be the same
CC or lower in the first, than in s2. The methods and products can be used
CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
CC or carcinogenicity of a test agent. AAX86201-33 represent SAGE tags
CC used to identify transcripts which are enhanced by p53.
XX
SQ Sequence 10 BP; 0 A; 6 C; 1 G; 3 T; 0 other;
OY
Query Match 100.0%; Score 6; DB 20; Length 10;
Best Local Similarity 66.7%; Pred. NO. 2.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 GCCGCU 6
11111
4 GCCTCT 9

```

Search completed: May 21, 2003, 04:50:11
 Job time : 157 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:37:42 ; Search time 35.25 Seconds
(without alignments)
52.200 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccucu 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/CTDUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 6 | 100.0 | 8 | 3 | US-08-859-954-49 |
| 2 | 6 | 100.0 | 10 | 3 | US-08-388-353-459 |
| 3 | 6 | 100.0 | 10 | 3 | US-08-388-353-460 |
| 4 | 6 | 100.0 | 10 | 3 | US-08-388-353-461 |
| 5 | 6 | 100.0 | 10 | 3 | US-08-388-353-462 |
| 6 | 6 | 100.0 | 10 | 3 | US-08-388-353-463 |
| 7 | 6 | 100.0 | 10 | 3 | US-08-488-551B-459 |
| 8 | 6 | 100.0 | 10 | 3 | US-08-488-551B-460 |
| 9 | 6 | 100.0 | 10 | 3 | US-08-488-551B-461 |
| 10 | 6 | 100.0 | 10 | 3 | US-08-488-551B-462 |
| 11 | 6 | 100.0 | 10 | 3 | US-08-488-551B-463 |
| 12 | 6 | 100.0 | 10 | 4 | US-09-235-614-2 |
| 13 | 6 | 100.0 | 10 | 4 | US-09-235-614-3 |
| 14 | 6 | 100.0 | 10 | 4 | US-09-235-614-4 |
| 15 | 6 | 100.0 | 10 | 4 | US-09-235-614-5 |
| 16 | 6 | 100.0 | 10 | 4 | US-09-235-614-6 |
| 17 | 6 | 100.0 | 10 | 4 | US-09-154-750A-4 |
| 18 | 6 | 100.0 | 12 | 2 | US-08-494-301A-16 |
| 19 | 6 | 100.0 | 12 | 2 | US-08-480-020B-18 |
| 20 | 6 | 100.0 | 12 | 2 | US-08-480-020B-21 |
| 21 | 6 | 100.0 | 12 | 2 | US-08-910-618-18 |
| 22 | 6 | 100.0 | 12 | 2 | US-08-910-618-21 |
| 23 | 6 | 100.0 | 12 | 3 | US-08-779-355-5 |
| 24 | 6 | 100.0 | 12 | 3 | US-08-671-824-19 |
| 25 | 6 | 100.0 | 12 | 3 | US-08-874-825-88 |
| 26 | 6 | 100.0 | 12 | 3 | US-08-938-835A-5 |
| 27 | 6 | 100.0 | 12 | 3 | US-08-663-824-88 |

| | | | | | | |
|----|---|-------|----|---|--------------------|--------------------|
| 28 | 6 | 100.0 | 12 | 4 | US-09-243-335-1 | Sequence 1, App1 |
| 29 | 6 | 100.0 | 12 | 4 | US-08-910-322-18 | Sequence 18, App1 |
| 30 | 6 | 100.0 | 12 | 4 | US-08-910-322-21 | Sequence 21, App1 |
| 31 | 6 | 100.0 | 12 | 4 | US-09-281-418-65 | Sequence 65, App1 |
| 32 | 6 | 100.0 | 12 | 4 | US-08-484-939A-18 | Sequence 18, App1 |
| 33 | 6 | 100.0 | 12 | 4 | US-08-484-939A-21 | Sequence 21, App1 |
| 34 | 6 | 100.0 | 12 | 4 | US-09-043-303-88 | Sequence 48, App1 |
| 35 | 6 | 100.0 | 12 | 4 | US-09-231-303-88 | Sequence 88, App1 |
| 36 | 6 | 100.0 | 13 | 1 | US-08-284-746-5 | Sequence 5, App1 |
| 37 | 6 | 100.0 | 13 | 1 | US-08-284-746-16 | Sequence 16, App1 |
| 38 | 6 | 100.0 | 13 | 1 | US-08-050-073-162 | Sequence 162, App1 |
| 39 | 6 | 100.0 | 14 | 1 | US-08-303-004-23 | Sequence 23, App1 |
| 40 | 6 | 100.0 | 14 | 1 | US-08-050-073-99 | Sequence 99, App1 |
| 41 | 6 | 100.0 | 14 | 1 | US-08-050-073-161 | Sequence 161, App1 |
| 42 | 6 | 100.0 | 14 | 1 | US-08-192-941-25 | Sequence 25, App1 |
| 43 | 6 | 100.0 | 14 | 1 | US-08-171-718-119 | Sequence 119, App1 |
| 44 | 6 | 100.0 | 14 | 2 | US-08-173-489C-324 | Sequence 324, App1 |
| 45 | 6 | 100.0 | 14 | 2 | US-08-232-087A-6 | Sequence 6, App1 |

ALIGNMENTS

RESULT 1
US-08-859-954-49
; Sequence 49, Application US/08859954
; Patent No. 6083695
; GENERAL INFORMATION:
; APPLICANT: Hardin, Susan H.
; APPLICANT: Homayouni, Ramin
; APPLICANT: Hardin, Paul E.
; TITLE OF INVENTION: Design and Optimized Primer Library for
; TITLE OF INVENTION: Gene Sequencing and Method Thereof
; NUMBER OF SEQUENCES: 566
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,954
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,782
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5325
; TELEFAX: 713/651-5246
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "oligonucleotide"
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; US-08-859-954-49

Query Match 100.0%; Score 6; DB 3; Length 8;

Best Local Similarity 66.7%; Pred. No. 3.6e+07;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCCUCU 6
111:1
Db 3 GCCTCT 8

RESULT 2

US-08-388-353-459/C
; Sequence 459, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 459:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-388-353-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCUCU 6
111:1
Db 10 GCCTCT 5

RESULT 3

US-08-388-353-460/C
; Sequence 460, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCUCU 6
111:1
Db 9 GCCTCT 4

RESULT 4

US-08-388-353-461/C
; Sequence 461, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 5
US-08-388-353-462/c
Sequence 462, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 7 GCCTCT 2

RESULT 6
US-08-388-353-463/c
Sequence 463, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 6 GCCTCT 1

RESULT 7
US-08-488-551B-459/c
Sequence 459, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: McPhee, Dale A.
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 10 GCCTCT 5

RESULT 8
US-08-488-551B-460/C
Sequence 460, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 9 GCCTCT 4

RESULT 9
US-08-488-551B-461/C
Sequence 461, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 10
US-08-488-551B-462/C
Sequence 462, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
|||:|:
DB 7 GCCTCT 2

RESULT 11
US-08-488-551B-463/C
Sequence 463, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
DB 6 GCCTCT 1

RESULT 12
US-09-235-614-2/C
Sequence 2, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.

```

: APPLICANT: CLARK, CHRISTOPHER L.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 2
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-2

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|:
Db 6 GCCTCT 1

RESULT 13
US-09-235-614-3/c
: Sequence 3, Application US/09235614
: Patent No. 6183966
: GENERAL INFORMATION:
: APPLICANT: GRAY, DONALD M.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 3
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-3

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|:
Db 7 GCCTCT 2

RESULT 14
US-09-235-614-4/c
: Sequence 4, Application US/09235614
: Patent No. 6183966
: GENERAL INFORMATION:
: APPLICANT: GRAY, DONALD M.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
```

```

: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 4
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|:
Db 8 GCCTCT 3

RESULT 15
US-09-235-614-5/c
: Sequence 5, Application US/09235614
: Patent No. 6183966
: GENERAL INFORMATION:
: APPLICANT: GRAY, DONALD M.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: TITLE OF INVENTION: SEQUENCES FOR ANTISENSE TARGETING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 5
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-5

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.1e+04;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|:
Db 9 GCCTCT 4
```

Search completed: May 21, 2003, 06:28:35
Job time : 35.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:39:53 : Search time 238 Seconds
(without alignments)
33.289 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6
Sequence: 1 gccucu 6

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapept 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| C 1 | 6 | 100.0 | 9 | US-09-990-186-588 | Sequence 588, App |
| C 2 | 6 | 100.0 | 9 | US-09-990-186-2084 | Sequence 2084, App |
| C 3 | 6 | 100.0 | 9 | US-09-990-186-2086 | Sequence 2086, App |
| C 4 | 6 | 100.0 | 9 | US-09-990-186-2219 | Sequence 2219, App |
| C 5 | 6 | 100.0 | 9 | US-09-990-186-2280 | Sequence 2280, App |
| C 6 | 6 | 100.0 | 9 | US-09-990-186-2471 | Sequence 2471, App |
| C 7 | 6 | 100.0 | 9 | US-09-990-186-2473 | Sequence 2473, App |
| C 8 | 6 | 100.0 | 9 | US-09-989-789-588 | Sequence 588, App |
| C 9 | 6 | 100.0 | 9 | US-09-989-789-2084 | Sequence 2084, App |
| C 10 | 6 | 100.0 | 9 | US-09-989-789-2086 | Sequence 2086, App |
| C 11 | 6 | 100.0 | 9 | US-09-989-789-2219 | Sequence 2219, App |
| C 12 | 6 | 100.0 | 9 | US-09-989-789-2280 | Sequence 2280, App |
| C 13 | 6 | 100.0 | 9 | US-09-989-789-2471 | Sequence 2471, App |
| C 14 | 6 | 100.0 | 9 | US-09-989-789-2473 | Sequence 2473, App |
| C 15 | 6 | 100.0 | 10 | US-09-990-186-617 | Sequence 617, App |
| C 16 | 6 | 100.0 | 10 | US-09-990-186-1279 | Sequence 1279, App |
| C 17 | 6 | 100.0 | 10 | US-09-990-186-1308 | Sequence 1308, App |
| C 18 | 6 | 100.0 | 10 | US-09-990-186-1313 | Sequence 1313, App |
| C 19 | 6 | 100.0 | 10 | US-09-989-789-617 | Sequence 617, App |

| | | | | | |
|------|---|-------|----|---------------------|--------------------|
| C 20 | 6 | 100.0 | 10 | US-09-989-789-1279 | Sequence 1279, App |
| C 21 | 6 | 100.0 | 10 | US-09-989-789-1308 | Sequence 1308, App |
| C 22 | 6 | 100.0 | 10 | US-09-989-789-1313 | Sequence 1313, App |
| C 23 | 6 | 100.0 | 10 | US-10-033-145-153 | Sequence 153, App |
| C 24 | 6 | 100.0 | 10 | US-10-033-145-174 | Sequence 174, App |
| C 25 | 6 | 100.0 | 10 | US-10-033-145-355 | Sequence 355, App |
| C 26 | 6 | 100.0 | 10 | US-10-033-145-566 | Sequence 566, App |
| C 27 | 6 | 100.0 | 10 | US-10-033-145-968 | Sequence 968, App |
| C 28 | 6 | 100.0 | 10 | US-10-033-145-1007 | Sequence 1007, App |
| C 29 | 6 | 100.0 | 10 | US-10-033-145-1120 | Sequence 1120, App |
| C 30 | 6 | 100.0 | 10 | US-10-033-145-1286 | Sequence 1286, App |
| C 31 | 6 | 100.0 | 10 | US-10-033-145-1609 | Sequence 1609, App |
| C 32 | 6 | 100.0 | 11 | US-09-249-155-83 | Sequence 25, App1 |
| C 33 | 6 | 100.0 | 11 | US-09-249-155-207 | Sequence 207, App |
| C 34 | 6 | 100.0 | 11 | US-09-249-155-239 | Sequence 239, App |
| C 35 | 6 | 100.0 | 11 | US-09-761-116-73 | Sequence 73, App1 |
| C 36 | 6 | 100.0 | 12 | US-09-384-472-21 | Sequence 18, App1 |
| C 37 | 6 | 100.0 | 12 | US-09-384-472-18 | Sequence 21, App1 |
| C 38 | 6 | 100.0 | 12 | US-08-591-4868-102 | Sequence 102, App |
| C 39 | 6 | 100.0 | 12 | US-09-978-600-198 | Sequence 198, App |
| C 40 | 6 | 100.0 | 14 | US-09-998-027-26 | Sequence 26, App1 |
| C 41 | 6 | 100.0 | 14 | US-09-504-231A-1431 | Sequence 1431, App |
| C 42 | 6 | 100.0 | 14 | US-09-274-553D-1431 | Sequence 1431, App |
| C 43 | 6 | 100.0 | 14 | US-10-056-414-119 | Sequence 119, App |
| C 44 | 6 | 100.0 | 15 | | |
| C 45 | 6 | 100.0 | 15 | | |

ALIGNMENTS

RESULT 1
US-09-990-186-588/c
; Sequence 588, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIT, Olanig
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-588

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||||
Db 8 GCCCYCT 3

RESULT 2
US-09-990-186-2084/c
; Sequence 2084, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIT, Olanig
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20

NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2084
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: example target
OTHER INFORMATION: DNA
US-09-990-186-2084

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 3
US-09-990-186-2086/c
Sequence 2086, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2086
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2086

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 4
US-09-990-186-2219/c
Sequence 2219, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2219
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2219

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 5
US-09-990-186-2280/c
Sequence 2280, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2280
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2280

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 6
US-09-990-186-2471/c
Sequence 2471, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2471
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2471

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
|||:|:
DB 8 GCCTCT 3

RESULT 7
US-09-990-186-2473/c

```
; Sequence 2473, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-2473

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
   8 GCCTCT 3

Db

RESULT 8
US-09-989-789-588/c
; Sequence 588, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-588

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
   8 GCCTCT 3

Db

RESULT 9
US-09-989-789-2084/c
; Sequence 2084, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
```

```
; SEQ ID NO 2084
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2084

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
   8 GCCTCT 3

Db

RESULT 10
US-09-989-789-2086/c
; Sequence 2086, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2086
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2086

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
   |||:|
   8 GCCTCT 3

Db

RESULT 11
US-09-989-789-2219/c
; Sequence 2219, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2219
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2219

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 12
US-09-989-789-2280/c
; Sequence 2280, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2280
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2280

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 13
US-09-989-789-2471/c
; Sequence 2471, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2471
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2471

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 14
US-09-989-789-2473/c
; Sequence 2473, Application US/09989789
; Patent No. US20020063379A1

; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2473

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+08;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

RESULT 15
US-09-990-186-617/c
; Sequence 617, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-617

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
|||:|:
Db 8 GCCTCT 3

Search completed: May 21, 2003, 06:44:44
Job time : 239 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:13:38 ; Search time 1094 Seconds
(without alignments)
88.824 Million cell updates/sec

Title: US-09-936-146-3

Perfect score: 6

Sequence: 1 gccucu 6

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 6 | 100.0 | 15 | 14 | BO511821 EST619236 |
| 2 | 6 | 100.0 | 16 | 14 | BQ789979 hage0036A |
| 3 | 6 | 100.0 | 19 | 9 | AI569191 tr83fl12.x |
| 4 | 6 | 100.0 | 19 | 17 | AZ429998 IM0214F16 |
| 5 | 6 | 100.0 | 19 | 17 | AZ475079 IM0293817 |
| 6 | 6 | 100.0 | 19 | 17 | AZ480905 IM0302N22 |

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|------------------------------|
| 7 | 6 | 100.0 | 19 | 17 | AZ509071 IM0351A21 |
| 8 | 6 | 100.0 | 19 | 17 | AZ651803 IM0522N11 |
| 9 | 6 | 100.0 | 19 | 17 | AZ659603 IM0537N06 |
| 10 | 6 | 100.0 | 19 | 17 | AZ800646 2M0058H14 |
| 11 | 6 | 100.0 | 19 | 17 | AZ834038 2M0116H01 |
| 12 | 6 | 100.0 | 20 | 14 | BQ789776 hage002aa |
| 13 | 6 | 100.0 | 20 | 17 | AZ309156 IM0013B09 |
| 14 | 6 | 100.0 | 20 | 17 | AZ366535 IM0115M15 |
| 15 | 6 | 100.0 | 20 | 17 | AZ615164 IM0444B24 |
| 16 | 6 | 100.0 | 20 | 17 | AZ797468 2M0053P09 |
| 17 | 6 | 100.0 | 20 | 17 | AZ807038 2M0069C06 |
| 18 | 6 | 100.0 | 21 | 17 | AZ303734 IM0003021 |
| 19 | 6 | 100.0 | 21 | 17 | AZ317085 IM0035D07 |
| 20 | 6 | 100.0 | 21 | 17 | AZ342282 IM0075H14 |
| 21 | 6 | 100.0 | 21 | 17 | AZ650869 IM0521G14 |
| 22 | 6 | 100.0 | 22 | 9 | AI023917 OW69F05.s |
| 23 | 6 | 100.0 | 22 | 9 | AI128425 q667f09.x |
| 24 | 6 | 100.0 | 22 | 9 | AI660937 wF20B06.x |
| 25 | 6 | 100.0 | 22 | 17 | AZ787102 2M0033B03 |
| 26 | 6 | 100.0 | 22 | 17 | AZ792724 2M0045F20 |
| 27 | 6 | 100.0 | 22 | 17 | AZ803482 2M0063I23 |
| 28 | 6 | 100.0 | 22 | 17 | TA204A05P AL478453 T. Brucei |
| 29 | 6 | 100.0 | 22 | 17 | TA294D03P AL485077 T. Brucei |
| 30 | 6 | 100.0 | 23 | 17 | AZ393604 IM0156C14 |
| 31 | 6 | 100.0 | 23 | 17 | AZ779967 2M0016C22 |
| 32 | 6 | 100.0 | 23 | 17 | AZ781980 2M0021N13 |
| 33 | 6 | 100.0 | 23 | 17 | AZ783504 2M0025L17 |
| 34 | 6 | 100.0 | 23 | 17 | AZ830526 2M0109M16 |
| 35 | 6 | 100.0 | 23 | 17 | AZ954682 2M0220P23 |
| 36 | 6 | 100.0 | 23 | 17 | AZ967993 2M0240J07 |
| 37 | 6 | 100.0 | 23 | 17 | BH812514 SALK_0618 |
| 38 | 6 | 100.0 | 24 | 17 | AZ443047 IM0237I06 |
| 39 | 6 | 100.0 | 24 | 17 | AZ463313 IM0272L03 |
| 40 | 6 | 100.0 | 24 | 17 | AZ505865 IM0346C18 |
| 41 | 6 | 100.0 | 24 | 17 | AZ780307 2M0017J04 |
| 42 | 6 | 100.0 | 24 | 17 | AZ787390 2M0033J23 |
| 43 | 6 | 100.0 | 24 | 17 | AZ789936 2M0038L17 |
| 44 | 6 | 100.0 | 24 | 17 | AZ812679 2M0079H03 |
| 45 | 6 | 100.0 | 24 | 17 | AZ936903 2M0193E20 |

ALIGNMENTS

RESULT 1
BO511821
LOCUS
DEFINITION
EST619236 generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone SMH018 5' end, mRNA sequence.

ACCESSION
BO511821
VERSION
BO511821.1 GI:21370690
KEYWORDS
SOURCE
ORGANISM

Solanum tuberosum

potato.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 15)

Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karamycheva, S.A.

Generation of a set of potato cDNA clones for microarray analyses

Unpublished (2002)

Other ESTs: EST619237

Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato@tigr.org

This clone is available through the Research Genetics, contact the

Research Genetics for further information 1-800-711-6195 or

cdna@resgen.com

Seq primer: T3.

FEATURES
source
Location/Qualifiers
1. 15
/organism="Solanum tuberosum"
/cultivar="Kennebec or Bintje"
/db_xref="taxon:4113"
/clone="STM018"
/clone_1lb="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
, tubers, or roots."

BASE COUNT
ORIGIN
3 a 5 c 2 g 5 t

Query Match 100.0%; Score 6; DB 14; Length 15;
Best Local Similarity 66.7%; Pred. NO. 5.5e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
Db 9 GCCTCT 14

RESULT 2
LOCUS B0789979 16 bp mRNA linear EST 30-JUL-2002
DEFINITION hage005a12 Heterobasidion annosum - Scots pine infection stage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage005a12, mRNA sequence.
B0789979
B0789979.1 GI:22004941
EST.
Pinus sylvestris/Heterobasidion annosum.
Pinus sylvestris/Heterobasidion annosum
Eukaryota: mixed EST libraries.
1 (bases 1 to 16)
Asiegbu F.O., Nahalkova, J. and Dean, R. A.
Selected Expressed sequence tags of cDNA clones from the
interaction of the root rot fungus (Heterobasidion annosum) with
seedling roots of Scots pine (Pinus sylvestris)
Unpublished (2001)
Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala,
Sweden
Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer.

FEATURES
source
Location/Qualifiers
1. 16
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone="hage005a12"
/clone_1lb="Heterobasidion annosum - Scots pine infection
stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for
6 days with H. annosum"
/note="Vector: pT-Adv; Site_1: EcoRI; The subtractive
hybridization cDNA library was constructed from scots pine
roots infected for 6 days with mycelia of Heterobasidion
annosum (FP5)."

BASE COUNT
ORIGIN
5 a 3 c 3 g 5 t

Query Match 100.0%; Score 6; DB 14; Length 16;
Best Local Similarity 66.7%; Pred. NO. 5.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
Db 7 GCCTCT 12

RESULT 3
LOCUS A1569191 19 bp mRNA linear EST 14-MAY-1999
DEFINITION tr83f12.x1 NCI-CGAP_Pan1 Homo sapiens cDNA clone IMAGE:224943 3'
similar to SW:PCEC_HUMAN P28074 PROTEASOME EPSILON CHAIN PRECURSOR
, mRNA sequence.

ACCESSION A1569191
VERSION A1569191.1 GI:4532565
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi;
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 19)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Trace considered overall poor quality
Insert Length: 1117 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1
POLA-No.

FEATURES
source
Location/Qualifiers
1. 19
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:224943"
/clone_1lb="NCI-CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: PCMV-SPORE; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT
ORIGIN
2 a 7 c 5 g 5 t

Query Match 100.0%; Score 6; DB 9; Length 19;
Best Local Similarity 66.7%; Pred. NO. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCUCU 6
Db 8 GCCTCT 13

RESULT 4
LOCUS A2429998 19 bp DNA linear GSS 03-OCT-2000
DEFINITION IM0214F16F Mouse 10kb plasmid UNGC1M library Mus musculus genomic
clone UNGC1M0214F16 F, DNA sequence.

ACCESSION A2429998
VERSION A2429998
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0214 row: F column: 16
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0214P16"
/clone_1lb="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
5 a 3 c 9 g 2. t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:1

Db 11 GCCTCT 6

RESULT 5
AZ475079/c 19 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION
1M0293B17F Mouse 10kb plasmid UGGCM library Mus musculus genomic
clone UGGCM0293B17 F, DNA sequence.
ACCESSION
AZ475079
VERSION
AZ475079.1 GI:10633204
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0293 row: B column: 17
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0293B17"
/clone_1lb="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 4 c 9 g 0 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:1

Db 18 GCCTCT 13

RESULT 6
AZ480905 19 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION
1M0302N22R Mouse 10kb plasmid UGGCM library Mus musculus genomic
clone UGGCM0302N22 R, DNA sequence.
ACCESSION
AZ480905
VERSION
AZ480905.1 GI:10642066
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 22
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0302N22"
/clone_1lb="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|gblAP129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
1 a 5 c 9 g 4 t

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gccucb 6
||||:
Db 7 GCCTCT 12

RESULT 7
A2509071/c 19 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0351A21R Mouse 10kb plasmid UGC1M library Mus musculus genomic
DEFINITION clone UGC1M0351A21 R, DNA sequence.
ACCESSION A2509071 GI:10690387
VERSION A2509071.1
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0351 row: A column: 21
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0351A21"
/clone_1lb="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|gblAP129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
4 a 4 c 9 g 2 t

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gccucu 6
||||:
Db 13 GCCTCT 8

RESULT 8
A2651803 19 bp DNA linear GSS 14-DEC-2000
LOCUS 1M0522N11R Mouse 10kb plasmid UGC1M library Mus musculus genomic
DEFINITION clone UGC1M0522N11 R, DNA sequence.
ACCESSION A2651803
VERSION A2651803.1
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
.M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0522 row: N column: 11
Seq primer: CACACAGGAACGATATGACC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0522N11"
/clone_1lb="Mouse 10kb plasmid library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|9b|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 7 c 5 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:|

DB 7 GCCTCT 12

RESULT 9
AZ659603/c 19 bp DNA linear GSS 14-DEC-2000
LOCUS 1M0537N06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0537N06 F, DNA sequence.
ACCESSION AZ659603
VERSION AZ659603
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
.M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0537 row: N column: 06
Seq primer: CGTGTAAACGACGCCACT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0537N06"
/clone_1lb="Mouse 10kb plasmid library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|9b|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
7 a 2 c 7 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:|

DB 12 GCCTCT 7

RESULT 10
AZ800646/c 19 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0058H14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0058H14 R, DNA sequence.
ACCESSION AZ800646
VERSION AZ800646
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M.', Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0058 row: H column: 14
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0058H14"
/clone.lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (g1147321149b1AF129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance.

BASE COUNT
3 a 4 c 7 g 5 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:
||||:
Db 7 GCCTCT 2

RESULT 11
A2834038 19 bp DNA linear GSS 20-FEB-2001
LOCUS A2834038 Mouse 10kb plasmid UGCGM library Mus musculus genomic
DEFINITION clone UGCG2M0116H01 R, DNA sequence.
ACCESSION A2834038
VERSION A2834038.1 GI:13003946
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M.', Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0116 row: H column: 01
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0116H01"
/clone.lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (g1147321149b1AF129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance.

BASE COUNT
6 a 2 c 7 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 66.7%; Pred. No. 6.2e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
||||:
||||:
Db 6 GCCTCT 1

RESULT 12
BQ789776 20 bp mRNA linear EST 30-JUL-2002
LOCUS BQ789776
DEFINITION hage002a10 Heterobasidion annosum - Scots pine infection stage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage002a10, mRNA sequence.
ACCESSION BQ789776
VERSION BQ789776.1 GI:22004738
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 20)

AUTHORS Asiegbu, F.O., Nahalkova, J. and Dean, R.A.

TITLE Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (Heterobasidion annosum) with seedling roots of Scots pine (Pinus sylvestris)

JOURNAL Unpublished (2001)

COMMENT Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden

Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se

Seq primer: T7 primer.

FEATURES
source Location/Qualifiers

1. 20
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone_lib="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
/note="Vector: pT-Adv; Site: 1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of Heterobasidion annosum (FP5)."
BASE COUNT 4 a 5 c 3 g 7 t 1 others
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 6; DB 14; Length 20;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
DB 12 GCCTCT 17

RESULT 13

AZ309156

LOCUS 20 bp DNA 1linear GSS 29-SEP-2000

DEFINITION IM0013809F Mouse 10kb plasmid UGCG1M library Mus musculus genomic

ACCESSION AZ309156

VERSION AZ309156.1 GI:10349862

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0013 row: B column: 09

Seq primer: CGTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 20.

FEATURES
source Location/Qualifiers

1. 20

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UGCG1M0013809"

/clone_lib="Mouse 10kb plasmid UGCG1M library"

/sex="Male"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gil147321419b/AP129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 2 a 11 c 4 g 3 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 6; DB 17; Length 20;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCUCU 6
DB 3 GCCTCT 8

RESULT 14

AZ366535

LOCUS 20 bp DNA 1linear GSS 02-OCT-2000

DEFINITION IM0115M15R Mouse 10kb plasmid UGCG1M library Mus musculus genomic

ACCESSION AZ366535

VERSION AZ366535.1 GI:10480235

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0115 row: M column: 15

Seq primer: CACACAGGAACGACTATACAC

Class: plasmid ends

High quality sequence stop: 20.

FEATURES
source Location/Qualifiers

1. 20

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UGGCM0115M15"
/clone_1lb="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      5 a      6 c      3 g      6 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GCCUCU 6
      |||:|:
      14 GCCTCT 19

```

```

RESULT 15
A2615164
LOCUS      20 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION clone UGGCM0444B24 F, DNA sequence.
ACCESSION  A2615164
VERSION    A2615164.1 GI:11737354
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE  1 (bases 1 to 20)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)

```

```

TITLE      University of Utah
JOURNAL    University of Utah
COMMENT    Contact: Robert B. Weiss
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177

```

```

FEATURES
Source     Email: dunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0444 Row: B Column: 24
            Seq primer: CGTTGTAACGACGCCACT
            Class: plasmid ends
            High quality sequence stop: 20.
            Location/Qualifiers
            1..20

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UGGCM0444B24"
/clone_1lb="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      3 a      7 c      5 g      5 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GCCUCU 6
      |||:|:
      DB      7 GCCTCT 12

```

```

Search completed: May 21, 2003, 06:26:03
Job time : 1095 secs

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 05:13:24 ; Search time 331.5 seconds

(without alignments)
526.748 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|-------------------|
| C 1 | 6 | 100.0 | 10 | 6 | AR128987 | AR128987 Sequence |
| C 2 | 6 | 100.0 | 10 | 6 | AR128988 | AR128988 Sequence |
| C 3 | 6 | 100.0 | 10 | 6 | AR128989 | AR128989 Sequence |
| C 4 | 6 | 100.0 | 10 | 6 | AR128990 | AR128990 Sequence |
| C 5 | 6 | 100.0 | 10 | 6 | AR128991 | AR128991 Sequence |
| C 6 | 6 | 100.0 | 10 | 6 | AR152349 | AR152349 Sequence |
| C 7 | 6 | 100.0 | 10 | 6 | AR152349 | AR152349 Sequence |
| C 8 | 6 | 100.0 | 10 | 6 | AR152849 | AR152849 Sequence |
| C 9 | 6 | 100.0 | 10 | 6 | AR152886 | AR152886 Sequence |
| C 10 | 6 | 100.0 | 10 | 6 | AR153384 | AR153384 Sequence |
| C 11 | 6 | 100.0 | 10 | 6 | AR153513 | AR153513 Sequence |
| C 12 | 6 | 100.0 | 10 | 6 | AR153513 | AR153513 Sequence |
| C 13 | 6 | 100.0 | 11 | 6 | AR1501 | AR1501 Sequence |
| C 14 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 15 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 16 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 17 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 18 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 19 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 20 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 21 | 6 | 100.0 | 11 | 6 | AR1503 | AR1503 Sequence |
| C 22 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 23 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 24 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 25 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 26 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 27 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 28 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 29 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 30 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 31 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 32 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 33 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 34 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 35 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 36 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 37 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 38 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 39 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 40 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 41 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 42 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 43 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 44 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |
| C 45 | 6 | 100.0 | 12 | 6 | AR1503 | AR1503 Sequence |

ALIGNMENTS

RESULT 1
LOCUS AR128987/c 10 bp DNA 1 linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6183966.
ACCESSION AR128987
VERSION AR128987.1 GI:14116649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURNAL Patent: US 6183966-A 2 06-FEB-2001.

FEATURES Location/Qualifiers
Source 1..10
BASE COUNT 2 a 3 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 6 GCCTCT 1

RESULT 2
LOCUS AR128988 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6183966.
ACCESSION AR128988
VERSION AR128988.1 GI:14116650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 3 06-FEB-2001;
FEATURES Location/Qualifiers
Source 1..10
BASE COUNT 2 a 3 c 5 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 7 GCCTCT 2

RESULT 3
LOCUS AR128989 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 4 from patent US 6183966.
ACCESSION AR128989
VERSION AR128989.1 GI:14116651
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 4 06-FEB-2001;
FEATURES Location/Qualifiers
Source 1..10
BASE COUNT 3 a 3 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 8 GCCTCT 3

RESULT 4
LOCUS AR128990 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6183966.
ACCESSION AR128990
VERSION AR128990.1 GI:14116652
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 5 06-FEB-2001;
FEATURES Location/Qualifiers
Source 1..10
BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 9 GCCTCT 4

RESULT 5
LOCUS AR128991 10 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6183966.
ACCESSION AR128991
VERSION AR128991.1 GI:14116653
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for
antisense targeting
JOURNAL Patent: US 6183966-A 6 06-FEB-2001;
FEATURES Location/Qualifiers
Source 1..10
BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 10 GCCTCT 5

RESULT 6
LOCUS AX152349 10 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 264 from Patent W00138577.
ACCESSION AX152349
VERSION AX152349.1 GI:14534000
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 264 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1.10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 0 a 5 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 1 GCCTCT 6

RESULT 7
AX152410 10 bp DNA PAT 22-JUN-2001
LOCUS AX152410
DEFINITION Sequence 325 from Patent WO0138577.
ACCESSION AX152410
VERSION AX152410.1 GI:14534061
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 325 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1.10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1 a 5 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 1 GCCTCT 6

RESULT 8
AX152849 10 bp DNA PAT 22-JUN-2001
LOCUS AX152849
DEFINITION Sequence 764 from Patent WO0138577.
ACCESSION AX152849
VERSION AX152849.1 GI:14534500
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 764 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1.10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1 a 3 c 3 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 5 GCCTCT 10

RESULT 9
AX152886 10 bp DNA PAT 22-JUN-2001
LOCUS AX152886
DEFINITION Sequence 801 from Patent WO0138577.
ACCESSION AX152886
VERSION AX152886.1 GI:14534537
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 801 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1.10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 5 a 1 c 3 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 10
AX153384 10 bp DNA PAT 22-JUN-2001
LOCUS AX153384
DEFINITION Sequence 1299 from Patent WO0138577.
ACCESSION AX153384
VERSION AX153384.1 GI:14535035
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 1299 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1.10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

```

Db      2 GCCTCT 7

RESULT 11
LOCUS   AX153513          10 bp   DNA      linear   PAT 22-JUN-2001
DEFINITION Sequence 1428 from Patent WO0138577.
ACCESSION AX153513
VERSION   AX153513.1 GI:14535164
KEYWORDS
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE    Human transcriptions
JOURNAL Patent: WO 0138577-A 1428 31-MAY-2001;
          The Johns Hopkins University (US)
FEATURES
source   1. .10
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
BASE COUNT      0 a      4 c      2 g      4 t
ORIGIN
Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 GCCTCU 6
        |||||:
Db      1 GCCTCT 6

RESULT 12
LOCUS   E16890          10 bp   DNA      linear   PAT 28-JUL-1999
DEFINITION DNA sequence required for efficient protein transcription in
          Brevibacterium flavum.
ACCESSION E16890
VERSION   E16890.1 GI:5711573
KEYWORDS JP 1998229881-A/31.
          Corynebacterium glutamicum.
          Corynebacterium glutamicum
          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
          Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Kobayashi,M., Man,T. and Yugawa,H.
TITLE    DNA HAVING SEQUENCE CAPABLE OF EFFICIENTLY TRANSLATING PROTEIN IN
          CORYNEFORM BACTERIA
JOURNAL Patent: JP 1998229881-A 31 02-SEP-1998;
          MITSUBISHI CHEM CORP
COMMENT   OS Brevibacterium flavum
          PN JP 1998229881-A/31
          PD 02-SEP-1998
          PI 19-FEB-1997 JP 1997035338
          PC C12N15/09,C07H21/04,C12N1/21//C12N9/38,C12Q1/68,(C12N15/09, PC
          C12R1:19),
          PC (C12N1/21,C12R1:13),(C12N9/38,C12R1:19);
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          FH key Location/Qualifiers
          FT source 1. .10
          /organism="Brevibacterium flavum" FT
          Location/Qualifiers
          1. .10
          /organism="Corynebacterium glutamicum"
FEATURES
source

```

```

BASE COUNT      3 a      1 c      5 g      1 t
ORIGIN
Query Match      100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 GCCTCU 6
        |||||:
Db      9 GCCTCT 4

RESULT 13
LOCUS   A91501          11 bp   DNA      linear   PAT 22-JAN-2000
DEFINITION Sequence 28 from Patent WO9824928.
ACCESSION A91501
VERSION   A91501.1 GI:6740456
KEYWORDS
SOURCE   unidentified.
          ORGANISM unidentified.
          REFERENCE 1 (bases 1 to 11)
          AUTHORS Pallsgaard,N. and Hokland,P.
          TITLE    DETECTION OF CHROMOSOMAL ABNORMALITIES
          JOURNAL Patent: WO 9824928-A 28 11-JUN-1998;
          PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
          Location/Qualifiers
          1. .11
          /organism="unidentified"
          /db_xref="taxon:32644"
BASE COUNT      1 a      5 c      2 g      3 t
ORIGIN
Query Match      100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2.7e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 GCCTCU 6
        |||||:
Db      5 GCCTCT 10

RESULT 14
LOCUS   A91503          11 bp   DNA      linear   PAT 22-JAN-2000
DEFINITION Sequence 30 from Patent WO9824928.
ACCESSION A91503
VERSION   A91503.1 GI:6740458
KEYWORDS
SOURCE   unidentified.
          ORGANISM unidentified.
          REFERENCE 1 (bases 1 to 11)
          AUTHORS Pallsgaard,N. and Hokland,P.
          TITLE    DETECTION OF CHROMOSOMAL ABNORMALITIES
          JOURNAL Patent: WO 9824928-A 30 11-JUN-1998;
          PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
          Location/Qualifiers
          1. .11
          /organism="unidentified"
          /db_xref="taxon:32644"
BASE COUNT      0 a      5 c      2 g      4 t
ORIGIN
Query Match      100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2.7e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 GCCTCU 6
        |||||:
Db      5 GCCTCT 10

```

RESULT 15
AX470495/c 11 bp DNA linear PAT 09-AUG-2002
LOCUS AX470495
DEFINITION Sequence 72 from Patent WO02053773.
ACCESSION AX470495
VERSION AX470495.1 GI:22205620
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Hofmann, K., Conradt, M. and Petersohn, D.
TITLE Method for determining skin stress or skin ageing in vitro
JOURNAL Patent: WO 02053773-A 72 11-JUL-2002;
HENKEL KGAA (DE)
FEATURES
source Location/Qualifiers
1..11
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 6 a 2 c 3 g 0 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred.No. 2.7e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3
Search completed: May 21, 2003, 07:17:19
Job time : 331.5 secs

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

ALIGNMENTS

```

RESULT 1
ID      AAX29578
XX      AAX29578 standard; DNA; 8 BP.
AC      AAX29578;
XX
XX      03-JUN-1999 (first entry)
DT
DE
XX      Primer for human G-protein coupled receptor genes.
XX
XX      Nucleic acid amplification; nuclear receptor; G-protein coupled receptor;
XX      apoptosis; DNA repair; DNA replication; plant biology; agriculture;
XX      human; veterinary medicine; reproduction; microbiology; hybridisation;
XX      environmental science; DNA fingerprinting; PCR primer; ss.
XX
XX      Synthetic.
XX      OS
XX      Homo sapiens.
XX      OS
XX      WO9911823-A2.
XX      PN
XX      11-MAR-1999.
XX      PD
XX      04-SEP-1998; 98WO-US18392.
XX      PF
XX      05-SEP-1997; 97US-0925816.
XX      PR
XX      (KIMM-) KIMMEL CANCER CENT SIDNEY.
XX      PA
XX      McCelland M, Pesole G;
XX      PI
XX      WPI; 1999-205200/17.
XX      DR
XX
```

PT Subset of primers able to amplify group of related sequences
 XX
 PS Claim 21; Page 75; 92pp; English.
 XX
 CC The invention provides primers (AAx29501-X29679) for identifying
 CC sequences encoding structurally or functionally related proteins such as
 CC nuclear or G-protein coupled receptors, apoptosis-related or DNA
 CC repair/replication proteins. The identified sequences are broadly useful
 CC in plant biology, agriculture, human or veterinary medicine,
 CC reproduction, microbiology or environmental science, e.g. to study
 CC expression of nuclear receptors at different stages of tissue development
 CC or after treatment with particular drugs. It is also used for DNA
 CC fingerprinting (to generate products useful for differential
 CC hybridisation), or, where a 3'-anchor primer is used, to isolate the
 CC 3'-ends of mRNA sequences. Sequences AAx29526-X29583 represent claimed
 CC primers specific for human G-protein coupled receptor genes.
 CC
 SO Sequence 8 BP; 0 A; 4 C; 2 G; 2 T; 0 other;
 Query Match 100.0%; Score 6; DB 20; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.6e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCT 6
 Db 1 GCCTCT 6
 RESULT 2
 ID AAA80736 standard; DNA; 8 BP.
 AC AAA80736;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE A. thaliana primer walking octamer SEQ ID NO: 49.
 KW
 XX Primer walking; octamer; primer; DNA sequencing; PCR; ss.
 OS Arabidopsis thaliana.
 XX
 PN US6083695-A.
 PD 04-JUL-2000.
 XX
 PF 21-MAY-1997; 97US-0859954.
 XX
 PR 15-APR-1996; 96US-0632782.
 XX
 PA (UW-HO-) UNIV HOUSTON.
 PA (HARD/) HARDIN S H.
 PI Hardin PE, Hardin SH, Homayouni R;
 DR WPI: 2000-474852/41.
 XX
 PT Sequencing an unknown DNA molecule for the polymerase chain reaction
 PT and other primer processes comprises primer walking of octamer
 PT oligonucleotides -
 PS
 XX Example 8; Column 49-50; 161pp; English.
 CC This invention describes a novel method for sequencing an unknown DNA
 CC molecule which comprises selecting a library primer from an octamer
 CC oligonucleotide library consisting of 48 8-bp sequences and
 CC corresponding complementary sequences, where the library primer is
 CC complementary to a known sequence adjacent to the unknown sequence or
 CC is complementary to a sequence in a known extension product. The method
 CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
 CC which make use of primers. The octamers are used to identify coding
 CC sequences. Primer walking using the octamer libraries is advantageous
 CC over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
 CC directed and methodical approach. AAA80688-A81253 represent the octamer
 CC primers used in the primer walking method of the invention.
 XX
 SO Sequence 8 BP; 0 A; 3 C; 1 G; 4 T; 0 other;
 Query Match 100.0%; Score 6; DB 21; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.6e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCT 6
 Db 3 GCCTCT 8
 RESULT 3
 ID ABQ71469/C standard; DNA; 9 BP.
 AC ABQ71469;
 XX
 DT 28-AUG-2002 (first entry)
 XX
 DE Zinc finger protein related oligonucleotide target SEQ ID NO:588.
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200242459-A2.
 PD 30-MAY-2002.
 XX
 PE 20-NOV-2001; 2001WO-US43438.
 XX
 PR 20-NOV-2000; 2000US-0716637.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 PA Liu Q;
 PI WPI: 2002-500284/53.
 DR
 XX
 PT New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 PS
 XX Example 1; Page 44; 81pp; English.
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3',-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determined the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABQ48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.


```
XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;
SQ Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.3e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
Db 8 GCCTCT 3

RESULT 4
ABQ71786/c
ID ABQ71786 standard; DNA: 9 BP.
XX ABQ71786;
XX
XX 28-AUG-2002 (first entry)
XX
XX zinc finger protein related oligonucleotide target SEQ ID NO:2084.
DE
XX zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX Homo sapiens.
OS
XX Synthetic.
XX
XX MO200242459-A2.
XX
XX 30-MAY-2002.
XX
XX 20-NOV-2001; 2001MO-US43438.
XX
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Liu Q;
XX
XX WPI; 2002-500284/53.
XX
XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering,
XX comprises first, second and third zinc fingers, ordered from N- to
XX C-terminus
XX
XX Example 1; Page 55; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target sub-site. Also described are: (1) a polypeptide
XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target sub-site, selecting the F2 zinc finger such
XX that it binds to the S2 target sub-site, and selecting the F3 zinc
XX finger such that it binds to the S3 target sub-site, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target sub-sites having the nucleotide G in the 5'-most position of the
XX sub-site. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.
XX
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;
```

```
XX Query Match 100.0%; Score 6; DB 24; Length 9;
XX Best Local Similarity 83.3%; Pred. No. 2.3e+08;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 GCCTCU 6
XX Db 8 GCCTCT 3

XX RESULT 5
XX ABQ71788/c
XX ID ABQ71788 standard; DNA: 9 BP.
XX
XX AC ABQ71788;
XX
XX 28-AUG-2002 (first entry)
XX
XX zinc finger protein related oligonucleotide target SEQ ID NO:2086.
DE
XX zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX Homo sapiens.
OS
XX Synthetic.
XX
XX MO200242459-A2.
XX
XX 30-MAY-2002.
XX
XX 20-NOV-2001; 2001MO-US43438.
XX
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Liu Q;
XX
XX WPI; 2002-500284/53.
XX
XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering,
XX comprises first, second and third zinc fingers, ordered from N- to
XX C-terminus
XX
XX Example 1; Page 55; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target sub-site. Also described are: (1) a polypeptide
XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target sub-site, selecting the F2 zinc finger such
XX that it binds to the S2 target sub-site, and selecting the F3 zinc
XX finger such that it binds to the S3 target sub-site, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target sub-sites having the nucleotide G in the 5'-most position of the
XX sub-site. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.
XX
XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;
```

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 6

AB071921/c
ID AB071921 standard; DNA: 9 BP.

XX AC AB071921;

XX DT 28-AUG-2002 (first entry)

XX DE Zinc finger protein related oligonucleotide target SEQ ID NO:2219.

XX KM Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN W0200242459-A2.

XX PD 30-MAY-2002.

XX PF 20-NOV-2001; 2001MO-US43438.

XX PR 20-NOV-2000; 2000US-0716637.

XX PA (SANG-) SANGAMO BIOSCIENCES INC.

XX PI Liu Q;

XX WP: 2002-500284/53.

XX PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -

XX PS Example 1; Page 58; 81bp; English.

XX CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3',-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX SQ Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.3e+08;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

Db 8 GCCTCT 3

RESULT 7

AB071982/c
ID AB071982 standard; DNA: 9 BP.

XX AC AB071982;

XX DT 28-AUG-2002 (first entry)

XX DE Zinc finger protein related oligonucleotide target SEQ ID NO:2280.

XX KM Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN W0200242459-A2.

XX PD 30-MAY-2002.

XX PF 20-NOV-2001; 2001MO-US43438.

XX PR 20-NOV-2000; 2000US-0716637.

XX PA (SANG-) SANGAMO BIOSCIENCES INC.

XX PI Liu Q;

XX WP: 2002-500284/53.

XX PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -

XX PS Example 1; Page 59; 81bp; English.

XX CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3',-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX SQ Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.3e+08;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

```
RESULT 8
ABQ72173/C
ID ABQ72173 standard; DNA: 9 BP.
XX
XX
AC ABQ72173;
XX
XX
DT 28-AUG-2002 (first entry)
XX
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2471.
XX
XX
KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX
OS Homo sapiens.
XX
XX
OS Synthetic.
XX
XX
PN MO200242459-A2.
XX
XX
PD 30-MAY-2002.
XX
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
XX
PR 20-NOV-2000; 2000US-0716637.
XX
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX
PI Liu Q;
XX
XX
DR WPI; 2002-500284/53.
XX
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
XX
PS Example 1; Page 63; 81pp; English.
XX
XX
CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ72123 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. NO. 2.3e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX
QY 1 GCCTCU 6
| | | | |
Db 8 GCCTCT 3
| | | | |
XX
XX
RESULT 9
ABQ72175/C
ID ABQ72175 standard; DNA: 9 BP.
XX
```

```
XX
XX
AC ABQ72175;
XX
XX
DT 28-AUG-2002 (first entry)
XX
XX
DE Zinc finger protein related oligonucleotide target SEQ ID NO:2473.
XX
XX
KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
XX
XX
OS Homo sapiens.
XX
XX
OS Synthetic.
XX
XX
PN MO200242459-A2.
XX
XX
PD 30-MAY-2002.
XX
XX
PF 20-NOV-2001; 2001WO-US43438.
XX
XX
PR 20-NOV-2000; 2000US-0716637.
XX
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX
PI Liu Q;
XX
XX
DR WPI; 2002-500284/53.
XX
XX
PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -
XX
XX
PS Example 1; Page 63; 81pp; English.
XX
XX
CC The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (I) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ72123 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
SQ Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
XX
XX
Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. NO. 2.3e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX
QY 1 GCCTCU 6
| | | | |
Db 8 GCCTCT 3
| | | | |
XX
XX
RESULT 10
AAQ96863/C
ID AAQ96863 standard; DNA: 10 BP.
XX
XX
AC AAQ96863;
XX
```

```

DT 26-MAR-1996 (first entry)
XX
XX HIV-1 NL4-3 nef gene nucleotide deletion 458.
DE
XX
XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX W09521912-A1.
PN
XX
XX 17-AUG-1995.
PD
XX
XX 14-FEB-1995; 95WO-AU00063.
PF
XX
XX 23-DEC-1994; 94AU-0000284.
PR 14-FEB-1994; 94AU-0003864.
PR 21-FEB-1994; 94AU-0004002.
XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI
XX WPI; 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
XX
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region: the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
XX Sequence 10 BP; 4 A; 1 C; 4 G; 1 T; 0 other:
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
Db 10 GCCTCT 5

```

```

XX
XX (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI
XX WPI; 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 194; 301pp; English.
XX
XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region: the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
XX Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other:
SQ
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
Db 9 GCCTCT 4

```

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene (AA096141). The resulting avirulent HIV strains are still capable of inducing an immune response in humans, and enable the generation of therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 13
AA096866/c
ID AA096866 standard; DNA; 10 BP.

XX AC AA096866;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 461.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PA (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

PS Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deancucleotides (AA096406-097018) from the nef gene and/or

CC 1 or more deancucleotides (AA097019-097166) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene

CC (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 5 A; 2 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 7 GCCTCT 2

RESULT 14

AA096867/c
ID AA096867 standard; DNA; 10 BP.

XX AC AA096867;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 462.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

XX 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PA (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

PS Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deancucleotides (AA096406-097018) from the nef gene and/or

CC 1 or more deancucleotides (AA097019-097166) from the LTR region; the

CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene

CC (AA096141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX Sequence 10 BP; 4 A; 2 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.9e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 6 GCCTCT 1

RESULT 15

AA096201
ID AA096201 standard; DNA; 10 BP.

XX AC AA096201;

DT 22-SEP-1999 (first entry)

XX SAGE tag used to identify transcripts which are enhanced by p53.

DE p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity;

KW neoplastic; p53 binding site; p53 promoter; SAGE tag; ss.

XX Homo sapiens.

XX WO9914356-A2.

XX 25-MAR-1999.

XX 17-SEP-1998; 98WO-US19300.
 PF
 XX
 PR 30-MAR-1998; 98US-0079817.
 PR 17-SEP-1997; 97US-0059153.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX
 PI Kinzler KW, Polyak K, Vogelstein B;
 XX
 DR WPT: 1999-443793/37.
 XX
 PT Use of p53 transcription tags to determine p53 status in, e.g.
 PT cancer diagnosis
 XX
 PS
 XX

Example 1: Page 25; 73pp; English.

CC The specification describes the use of p53 transcription tags for
 CC developing products to determine p53 status, to diagnose cancer
 CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
 CC A method for diagnosing cancer or determining p53 status in a sample
 CC suspected for being neoplastic comprises comparing the level of
 CC transcription of an RNA transcript in a first sample (s1) of a first
 CC tissue (t1) to the level of transcription of the transcript in a second
 CC sample (s2) of a second tissue (s2), where s1 is suspected of being
 CC neoplastic and s2 is a normal human tissue (of the same type) and the
 CC transcript is identified by a tag; and categorizing s1 as neoplastic
 CC or as having a mutant p53 when transcription is found to be the same
 CC or lower in the first, than in s2. The methods and products can be used
 CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
 CC or carcinogenicity of a test agent. AAX86201-33 represent SAGE tags
 CC used to identify transcripts which are enhanced by p53.
 XX

SQ Sequence 10 BP; 0 A; 6 C; 1 G; 3 T; 0 other;

Query Match 100.0%; Score 6; DB 20; Length 10;
 Best Local Similarity 83.3%; Pred. No. 2.9e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
 |||||:
 Db 4 GCCTCT 9

Search completed: May 21, 2003, 06:54:55
 Job time : 148.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:28:43 ; Search time 34.75 Seconds
(without alignments)
52.951 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6

Sequence: 1 gctctc 6

Scoring table: OLIGO_NTC

Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCFUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 6 | 100.0 | 8 | 3 | US-08-859-954-49 |
| 2 | 6 | 100.0 | 10 | 3 | US-08-388-353-459 |
| 3 | 6 | 100.0 | 10 | 3 | US-08-388-353-460 |
| 4 | 6 | 100.0 | 10 | 3 | US-08-388-353-461 |
| 5 | 6 | 100.0 | 10 | 3 | US-08-388-353-462 |
| 6 | 6 | 100.0 | 10 | 3 | US-08-388-353-463 |
| 7 | 6 | 100.0 | 10 | 3 | US-08-488-551B-459 |
| 8 | 6 | 100.0 | 10 | 3 | US-08-488-551B-460 |
| 9 | 6 | 100.0 | 10 | 3 | US-08-488-551B-461 |
| 10 | 6 | 100.0 | 10 | 3 | US-08-488-551B-462 |
| 11 | 6 | 100.0 | 10 | 3 | US-08-488-551B-463 |
| 12 | 6 | 100.0 | 10 | 4 | US-09-235-614-2 |
| 13 | 6 | 100.0 | 10 | 4 | US-09-235-614-3 |
| 14 | 6 | 100.0 | 10 | 4 | US-09-235-614-4 |
| 15 | 6 | 100.0 | 10 | 4 | US-09-235-614-5 |
| 16 | 6 | 100.0 | 10 | 4 | US-09-235-614-6 |
| 17 | 6 | 100.0 | 10 | 4 | US-09-154-750A-4 |
| 18 | 6 | 100.0 | 12 | 2 | US-08-494-301A-16 |
| 19 | 6 | 100.0 | 12 | 2 | US-08-480-020B-18 |
| 20 | 6 | 100.0 | 12 | 2 | US-08-480-020B-21 |
| 21 | 6 | 100.0 | 12 | 2 | US-08-910-618-18 |
| 22 | 6 | 100.0 | 12 | 2 | US-08-910-618-21 |
| 23 | 6 | 100.0 | 12 | 3 | US-08-779-355-5 |
| 24 | 6 | 100.0 | 12 | 3 | US-08-671-824-19 |
| 25 | 6 | 100.0 | 12 | 3 | US-08-874-825-88 |
| 26 | 6 | 100.0 | 12 | 3 | US-08-938-835A-5 |
| 27 | 6 | 100.0 | 12 | 3 | US-08-663-824-88 |

| | | | | | | |
|----|---|-------|----|---|--------------------|--------------------|
| 28 | 6 | 100.0 | 12 | 4 | US-09-243-335-1 | Sequence 1, Appl |
| 29 | 6 | 100.0 | 12 | 4 | US-08-910-322-18 | Sequence 18, Appl |
| 30 | 6 | 100.0 | 12 | 4 | US-08-910-322-21 | Sequence 21, Appl |
| 31 | 6 | 100.0 | 12 | 4 | US-09-281-418-65 | Sequence 65, Appl |
| 32 | 6 | 100.0 | 12 | 4 | US-08-484-939A-18 | Sequence 18, Appl |
| 33 | 6 | 100.0 | 12 | 4 | US-08-484-939A-21 | Sequence 21, Appl |
| 34 | 6 | 100.0 | 12 | 4 | US-09-043-149-48 | Sequence 48, Appl |
| 35 | 6 | 100.0 | 12 | 4 | US-09-231-303-88 | Sequence 88, Appl |
| 36 | 6 | 100.0 | 13 | 1 | US-08-284-746-5 | Sequence 5, Appl |
| 37 | 6 | 100.0 | 13 | 1 | US-08-284-746-16 | Sequence 16, Appl |
| 38 | 6 | 100.0 | 13 | 1 | US-08-050-073-162 | Sequence 162, Appl |
| 39 | 6 | 100.0 | 14 | 1 | US-08-303-004-23 | Sequence 23, Appl |
| 40 | 6 | 100.0 | 14 | 1 | US-08-050-073-99 | Sequence 99, Appl |
| 41 | 6 | 100.0 | 14 | 1 | US-08-050-073-161 | Sequence 161, Appl |
| 42 | 6 | 100.0 | 14 | 1 | US-08-192-941-25 | Sequence 25, Appl |
| 43 | 6 | 100.0 | 14 | 1 | US-08-171-718-119 | Sequence 119, Appl |
| 44 | 6 | 100.0 | 14 | 2 | US-08-173-489C-324 | Sequence 324, Appl |
| 45 | 6 | 100.0 | 14 | 2 | US-08-232-087A-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1
US-08-859-954-49
; Sequence 49, Application US/08859954
; Patent No. 6083695
; GENERAL INFORMATION:
; APPLICANT: Hardin, Susan H.
; APPLICANT: Homayouni, Ramin
; TITLE OF INVENTION: Design and Optimized Primer Library for
; TITLE OF INVENTION: Gene Sequencing and Method Thereof
; NUMBER OF SEQUENCES: 566
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,954
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,782
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5325
; TELEFAX: 713/651-5246
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Oligonucleotide"
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; US-08-859-954-49

Query Match 100.0%; Score 6; DB 3; Length 8;

Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
 |||||
Db 3 GCCTCT 8

RESULT 2
US-08-388-353-459/c
; Sequence 459, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 459:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-388-353-459
Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
 |||||
Db 10 GCCTCT 5

RESULT 3
US-08-388-353-460/c
; Sequence 460, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 460:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-388-353-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
 |||||
Db 9 GCCTCT 4

RESULT 4
US-08-388-353-461/c
; Sequence 461, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 5
US-08-388-353-462/C
Sequence 462, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTCU 6
|||||
Db 7 GCCTCT 2

RESULT 6
US-08-388-353-463/C
Sequence 463, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTCU 6
|||||
Db 6 GCCTCT 1

RESULT 7
US-08-488-551B-459/C
Sequence 459, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: McPhee, Dale A.
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGILIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 10 GCCTCT 5

RESULT 8
US-08-488-551B-460/C
Sequence 460, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGILIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 9 GCCTCT 4

RESULT 9
US-08-488-551B-461/C
Sequence 461, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGILIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
DB 8 GCCTCT 3

RESULT 10
US-08-488-551B-462/c
Sequence 462, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-462
Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
DB 7 GCCTCT 2

RESULT 11
US-08-488-551B-463/c
Sequence 463, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-463
Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
DB 6 GCCTCT 1
RESULT 12
US-09-235-614-2/c
Sequence 2, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.

```
; APPLICANT: CLARK, CHRISTOPHER L.
; TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
; FILE REFERENCE: 91556/66384
; CURRENT APPLICATION NUMBER: US/09/235,614
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/808,474
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 08/320,507
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-2
```

```
Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3,4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GCCTCU 6
    |||||
Db 6 GCCTCT 1
```

```
RESULT 13
US-09-235-614-3/c
; Sequence 3, Application US/09235614
; Patent No. 6183966
; GENERAL INFORMATION:
; APPLICANT: GRAY, DONALD M.
; APPLICANT: CLARK, CHRISTOPHER L.
; TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
; FILE REFERENCE: 91556/66384
; CURRENT APPLICATION NUMBER: US/09/235,614
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/808,474
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 08/320,507
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-3
```

```
Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3,4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GCCTCU 6
    |||||
Db 7 GCCTCT 2
```

```
RESULT 14
US-09-235-614-4/c
; Sequence 4, Application US/09235614
; Patent No. 6183966
; GENERAL INFORMATION:
; APPLICANT: GRAY, DONALD M.
; APPLICANT: CLARK, CHRISTOPHER L.
; TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
; FILE REFERENCE: 91556/66384
; CURRENT APPLICATION NUMBER: US/09/235,614
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/808,474
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 08/320,507
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4
```

```
; FILE REFERENCE: 91556/66384
; CURRENT APPLICATION NUMBER: US/09/235,614
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/808,474
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 08/320,507
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4
```

```
Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3,4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GCCTCU 6
    |||||
Db 8 GCCTCT 3
```

```
RESULT 15
US-09-235-614-5/c
; Sequence 5, Application US/09235614
; Patent No. 6183966
; GENERAL INFORMATION:
; APPLICANT: GRAY, DONALD M.
; APPLICANT: CLARK, CHRISTOPHER L.
; TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
; FILE REFERENCE: 91556/66384
; CURRENT APPLICATION NUMBER: US/09/235,614
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/808,474
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 08/320,507
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-5
```

```
Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3,4e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GCCTCU 6
    |||||
Db 9 GCCTCT 4
```

```
Search completed: May 21, 2003, 08:40:18
Job time : 34.75 secs
```

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:44:53 ; Search time 65 Seconds
(without alignments)
121.889 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 828747 seqs, 660231138 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCr_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1 | 6 | 100.0 | 9 | US-09-990-186-588 | Sequence 588, App |
| 2 | 6 | 100.0 | 9 | US-09-990-186-2084 | Sequence 2084, App |
| 3 | 6 | 100.0 | 9 | US-09-990-186-2086 | Sequence 2086, App |
| 4 | 6 | 100.0 | 9 | US-09-990-186-2219 | Sequence 2219, App |
| 5 | 6 | 100.0 | 9 | US-09-990-186-2280 | Sequence 2280, App |
| 6 | 6 | 100.0 | 9 | US-09-990-186-2471 | Sequence 2471, App |
| 7 | 6 | 100.0 | 9 | US-09-990-186-2473 | Sequence 2473, App |
| 8 | 6 | 100.0 | 9 | US-09-989-789-588 | Sequence 588, App |
| 9 | 6 | 100.0 | 9 | US-09-989-789-2084 | Sequence 2084, App |
| 10 | 6 | 100.0 | 9 | US-09-989-789-2086 | Sequence 2086, App |
| 11 | 6 | 100.0 | 9 | US-09-989-789-2219 | Sequence 2219, App |
| 12 | 6 | 100.0 | 9 | US-09-989-789-2280 | Sequence 2280, App |
| 13 | 6 | 100.0 | 9 | US-09-989-789-2471 | Sequence 2471, App |
| 14 | 6 | 100.0 | 9 | US-09-989-789-2473 | Sequence 2473, App |
| 15 | 6 | 100.0 | 10 | US-09-990-186-617 | Sequence 617, App |
| 16 | 6 | 100.0 | 10 | US-09-990-186-1179 | Sequence 1179, App |
| 17 | 6 | 100.0 | 10 | US-09-990-186-1308 | Sequence 1308, App |
| 18 | 6 | 100.0 | 10 | US-09-990-186-1313 | Sequence 1313, App |
| 19 | 6 | 100.0 | 10 | US-09-989-789-617 | Sequence 617, App |

| | | | | | |
|----|---|-------|----|---------------------|--------------------|
| 20 | 6 | 100.0 | 10 | US-09-989-789-1279 | Sequence 1279, App |
| 21 | 6 | 100.0 | 10 | US-09-989-789-1308 | Sequence 1308, App |
| 22 | 6 | 100.0 | 10 | US-09-989-789-1313 | Sequence 1313, App |
| 23 | 6 | 100.0 | 10 | US-10-033-145-153 | Sequence 153, App |
| 24 | 6 | 100.0 | 10 | US-10-033-145-174 | Sequence 174, App |
| 25 | 6 | 100.0 | 10 | US-10-033-145-355 | Sequence 355, App |
| 26 | 6 | 100.0 | 10 | US-10-033-145-566 | Sequence 566, App |
| 27 | 6 | 100.0 | 10 | US-10-033-145-968 | Sequence 968, App |
| 28 | 6 | 100.0 | 10 | US-10-033-145-1007 | Sequence 1007, App |
| 29 | 6 | 100.0 | 10 | US-10-033-145-1120 | Sequence 1120, App |
| 30 | 6 | 100.0 | 10 | US-10-033-145-1286 | Sequence 1286, App |
| 31 | 6 | 100.0 | 10 | US-10-033-145-1609 | Sequence 1609, App |
| 32 | 6 | 100.0 | 11 | US-09-249-155-25 | Sequence 25, App |
| 33 | 6 | 100.0 | 11 | US-09-249-155-83 | Sequence 83, App |
| 34 | 6 | 100.0 | 11 | US-09-249-155-207 | Sequence 207, App |
| 35 | 6 | 100.0 | 11 | US-09-249-155-239 | Sequence 239, App |
| 36 | 6 | 100.0 | 12 | US-10-131-591A-73 | Sequence 73, App |
| 37 | 6 | 100.0 | 12 | US-09-761-116-1 | Sequence 1, App |
| 38 | 6 | 100.0 | 12 | US-09-384-472-18 | Sequence 18, App |
| 39 | 6 | 100.0 | 12 | US-09-384-472-21 | Sequence 21, App |
| 40 | 6 | 100.0 | 14 | US-08-591-486B-102 | Sequence 102, App |
| 41 | 6 | 100.0 | 14 | US-09-978-600-198 | Sequence 198, App |
| 42 | 6 | 100.0 | 14 | US-09-998-027-26 | Sequence 26, App |
| 43 | 6 | 100.0 | 14 | US-09-504-231A-1431 | Sequence 1431, App |
| 44 | 6 | 100.0 | 14 | US-09-274-553D-1431 | Sequence 1431, App |
| 45 | 6 | 100.0 | 15 | US-10-056-414-119 | Sequence 119, App |

ALIGNMENTS

RESULT 1
US-09-990-186-588/c
; Sequence 588, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-588

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 2
US-09-990-186-2084/c
; Sequence 2084, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20

NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2084
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2084

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 3
US-09-990-186-2086/c
Sequence 2086, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990.186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2086
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2086

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 4
US-09-990-186-2219/c
Sequence 2219, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990.186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2219
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2219

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 5
US-09-990-186-2280/c
Sequence 2280, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990.186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2280
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2280

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 6
US-09-990-186-2471/c
Sequence 2471, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990.186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2471
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2471

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 7
US-09-990-186-2473/c

```
; Sequence 2473, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-2473

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
   |||||
Db 8 GCCTCT 3

RESULT 8
US-09-989-789-588/C
; Sequence 588, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-588

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
   |||||
Db 8 GCCTCT 3

RESULT 9
US-09-989-789-2084/C
; Sequence 2084, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
```

```
; SEQ ID NO 2084
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2084

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
   |||||
Db 8 GCCTCT 3

RESULT 10
US-09-989-789-2086/C
; Sequence 2086, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2086
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2086

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
   |||||
Db 8 GCCTCT 3

RESULT 11
US-09-989-789-2219/C
; Sequence 2219, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2219
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2219

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
   |||||
Db 8 GCCTCT 3
```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 12
US-09-989-789-2280/c
; Sequence 2280, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2280
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2280

Query Match
Best Local Similarity 83.3%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 13
US-09-989-789-2471/c
; Sequence 2471, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2471
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2471

Query Match
Best Local Similarity 83.3%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 14
US-09-989-789-2473/c
; Sequence 2473, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2473

Query Match
Best Local Similarity 83.3%; Score 6; DB 10; Length 9;
Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 15
US-09-990-186-617/c
; Sequence 617, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-617

Query Match
Best Local Similarity 83.3%; Score 6; DB 9; Length 10;
Pred. No. 2.5e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

Search completed: May 21, 2003, 08:44:52
Job time : 66 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:26:13 ; Search time 1092.5 Seconds
(without alignments)
88.945 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6

Sequence: 1 gccctc 6

Scoring table:

OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hlc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_ylt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 100.0 | 15 | 14 | BQ511821 | BQ511821 EST619236 |
| 2 | 100.0 | 16 | 14 | BQ788979 | BQ788979 hage005aa |
| 3 | 100.0 | 19 | 9 | A1569191 | A1569191 tr83f12.x |
| 4 | 100.0 | 19 | 17 | A2429998 | A2429998 IM0214F16 |
| 5 | 100.0 | 19 | 17 | A2475079 | A2475079 IM0293B17 |
| 6 | 100.0 | 19 | 17 | A2480905 | A2480905 IM0302N22 |

| | | | | | | | |
|---|----|---|-------|----|----|-----------|---------------------|
| c | 7 | 6 | 100.0 | 19 | 17 | A2509071 | A2509071 IM0351A21 |
| c | 8 | 6 | 100.0 | 19 | 17 | A2651803 | A2651803 IM0522N11 |
| c | 9 | 6 | 100.0 | 19 | 17 | A2659603 | A2659603 IM0537N06 |
| c | 10 | 6 | 100.0 | 19 | 17 | A2800646 | A2800646 2M0058H14 |
| c | 11 | 6 | 100.0 | 19 | 17 | A2834038 | A2834038 2M0116H01 |
| c | 12 | 6 | 100.0 | 20 | 14 | BQ789776 | BQ789776 hage002aa |
| c | 13 | 6 | 100.0 | 20 | 17 | A2309156 | A2309156 IM0013B09 |
| c | 14 | 6 | 100.0 | 20 | 17 | A2366535 | A2366535 IM0115M15 |
| c | 15 | 6 | 100.0 | 20 | 17 | A2615164 | A2615164 IM044AB24 |
| c | 16 | 6 | 100.0 | 20 | 17 | A2797468 | A2797468 2M0053P09 |
| c | 17 | 6 | 100.0 | 20 | 17 | A2807038 | A2807038 2M0069C06 |
| c | 18 | 6 | 100.0 | 21 | 17 | A2303734 | A2303734 IM0003021 |
| c | 19 | 6 | 100.0 | 21 | 17 | A2317085 | A2317085 IM0035D07 |
| c | 20 | 6 | 100.0 | 21 | 17 | A2342282 | A2342282 IM0075H14 |
| c | 21 | 6 | 100.0 | 21 | 17 | A2650869 | A2650869 IM0521G14 |
| c | 22 | 6 | 100.0 | 22 | 9 | A1023917 | A1023917 OM69F05.S |
| c | 23 | 6 | 100.0 | 22 | 9 | A1128425 | A1128425 qc67F09.x |
| c | 24 | 6 | 100.0 | 22 | 9 | A1660937 | A1660937 wF20B06.x |
| c | 25 | 6 | 100.0 | 22 | 17 | A2787102 | A2787102 2M0033B03 |
| c | 26 | 6 | 100.0 | 22 | 17 | A2792724 | A2792724 2M0045F20 |
| c | 27 | 6 | 100.0 | 22 | 17 | A2803482 | A2803482 2M0063I23 |
| c | 28 | 6 | 100.0 | 22 | 17 | TA204A05P | TA204A05P T. brucei |
| c | 29 | 6 | 100.0 | 22 | 17 | TA294D03P | TA294D03P T. brucei |
| c | 30 | 6 | 100.0 | 23 | 17 | A2393604 | A2393604 IM0156C14 |
| c | 31 | 6 | 100.0 | 23 | 17 | A2779967 | A2779967 2M0016C22 |
| c | 32 | 6 | 100.0 | 23 | 17 | A2781980 | A2781980 2M0021M13 |
| c | 33 | 6 | 100.0 | 23 | 17 | A2783504 | A2783504 2M0025L17 |
| c | 34 | 6 | 100.0 | 23 | 17 | A2830526 | A2830526 2M0109M16 |
| c | 35 | 6 | 100.0 | 23 | 17 | A2954682 | A2954682 2M0220P23 |
| c | 36 | 6 | 100.0 | 23 | 17 | A2967993 | A2967993 2M0240J07 |
| c | 37 | 6 | 100.0 | 23 | 17 | BH812514 | BH812514 SALK_0618 |
| c | 38 | 6 | 100.0 | 24 | 17 | A2443047 | A2443047 IM0237I06 |
| c | 39 | 6 | 100.0 | 24 | 17 | A2463313 | A2463313 IM0272L03 |
| c | 40 | 6 | 100.0 | 24 | 17 | A2505865 | A2505865 IM0346C18 |
| c | 41 | 6 | 100.0 | 24 | 17 | A2780307 | A2780307 2M0017J04 |
| c | 42 | 6 | 100.0 | 24 | 17 | A2787390 | A2787390 2M0033J23 |
| c | 43 | 6 | 100.0 | 24 | 17 | A2789936 | A2789936 2M0038L17 |
| c | 44 | 6 | 100.0 | 24 | 17 | A2812679 | A2812679 2M0079H03 |
| c | 45 | 6 | 100.0 | 24 | 17 | A2936903 | A2936903 2M0193E20 |

ALIGNMENTS

RESULT 1
BQ511821
LOCUS
DEFINITION
EST619236 Generation of a set of potato CDNA clones for microarray analyses mixed potato tissues Solanum tuberosum CDNA clone STMH018 5' end, mRNA sequence.

ACCESSION
BQ511821
VERSION
BQ511821.1 GI:21370690

KEYWORDS
SOURCE
ORGANISM

Solanum tuberosum
potato.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karamycheva, S.A.
Generation of a set of potato CDNA clones for microarray analyses

Unpublished (2002)
Other ESTs: EST619237
Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potatoclitgr.org
This clone is available through the Research Genetics, contact the Research Genetics for further information 1-800-711-6195 or cdna@resgen.com
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1. 15
/organism="Solanum tuberosum"
/cultivar="Kennebec or Blonje"
/db_xref="taxon:4113"
/clone="STM018"
/clone_1lb="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: Bluescript SK(-). Site_1: EcoRI; Site_2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes / tubers, or roots."

BASE COUNT
origin 3 a 5 c 2 g 5 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 14; Length 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
Db 9 GCCTCT 14

RESULT 2
LOCUS B0789979
DEFINITION B0789979 16 bp mRNA linear EST 30-JUL-2002
hage005a12 Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion annosum cDNA clone hage005a12, mRNA sequence.
B0789979
B0789979.1 GI:22004941
EST.
Pinus sylvestris/Heterobasidion annosum.
Eukaryota; mixed EST libraries.
1 (bases 1 to 16)
Aalegbu.F.O., Nahalkova.J. and Dean.R.A.
Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (Heterobasidion annosum) with seedling roots of Scots pine (Pinus sylvestris)
Unpublished (2001)
Contact: Fred O. Aalegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden
Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Aalegbu@mykopat.slu.se
Seq primer: T7 primer
Location/Qualifiers
1. 16
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone="hage005a12"
/clone_1lb="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
/note="Vector: PT-Adv; Site_1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of Heterobasidion annosum (fp5)."

BASE COUNT
origin 5 a 3 c 3 g 5 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 14; Length 16;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
Db 7 GCCTCT 12

RESULT 3
LOCUS A1569191
DEFINITION A1569191 19 bp mRNA linear EST 14-MAY-1999
tr83f12.x1 NCI-CGAP Pan1 Homo sapiens cDNA clone IMAGE:2224943
similar to SW:PRCE_HUMAN P28074 PROTEASOME EMISSION CHAIN PRECURSOR
; mRNA sequence.

ACCESSION A1569191
VERSION A1569191.1 GI:4532565
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 19)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/dbfp/image/image.html

Trace considered overall poor quality
Insert length: 1117 Std Error: 0.00
Seq primer: -40UP from G1bc
High quality sequence stop: 1
POLYA-NO.

FEATURES
source
Location/Qualifiers
1. 19
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2224943"
/clone_1lb="NCI-CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH108"
/note="Organ: pancreas; Vector: PCMV-SPORE6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt".
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT
origin 2 a 7 c 5 g 5 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
Db 8 GCCTCT 13

RESULT 4
LOCUS A2429998/c
DEFINITION A2429998 19 bp DNA linear GSS 03-OCT-2000
clone U06C1M0214F16 F, DNA sequence.
A2429998
A2429998.1 GI:10554011
GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0214 row: F column: 16
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U0GCI0214F16"
/clone_1lb="Mouse 10kb plasmid U0GCI01 library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
5 a 3 c 9 g 2 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
|||||

Db 11 GCCTCT 6

RESULT 5
A2475079 19 bp. DNA linear GSS 04-OCT-2000
LOCUS A2475079
DEFINITION 1M0233B17F Mouse 10kb plasmid U0GCI01 library Mus musculus genomic
clone U0GCI0233B17 F, DNA sequence.
ACCESSION A2475079
VERSION A2475079.1 GI:10633204
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0293 row: B column: 17
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U0GCI0293B17"
/clone_1lb="Mouse 10kb plasmid U0GCI01 library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 4 c 9 g 0 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
|||||

Db 18 GCCTCT 13

RESULT 6
A2480905 19 bp. DNA linear GSS 04-OCT-2000
LOCUS A2480905
DEFINITION 1M0302N22R Mouse 10kb plasmid U0GCI01 library Mus musculus genomic
clone U0GCI0302N22 R, DNA sequence.
ACCESSION A2480905
VERSION A2480905.1 GI:10642066
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 22
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0302N22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1147321149b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
1 a 5 c 9 g 4 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 17; Length 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 7 GCCTCT 12

RESULT 7
A2509071/c 19 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0351A21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0351A21 R, DNA sequence.
ACCESSION A2509071
VERSION A2509071.1 GI:10690387
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0351 row: A column: 21
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0351A21"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1147321149b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
4 a 4 c 9 g 2 t

Query Match
Best Local Similarity 100.0%; Score 6; DB 17; Length 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 13 GCCTCT 8

RESULT 8
A2651803 19 bp DNA linear GSS 14-DEC-2000
LOCUS 1M0522N11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0522N11 R, DNA sequence.
ACCESSION A2651803
VERSION A2651803.1 GI:11787672
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0522 row: N column: 11
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
source
location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0522N11"
/clone_lib="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 7 c 5 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred.No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 7 GCCTCT 12

RESULT 9
A2659603 19 bp DNA linear GSS 14-DEC-2000
LOCUS 1M0537N06F Mouse 10kb plasmid UGC1M library Mus musculus genomic
DEFINITION clone UGC1M0537N06 F, DNA sequence.
ACCESSION A2659603
VERSION A2659603.1 GI:11796749
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0537 row: N column: 06
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
source
location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0537N06"
/clone_lib="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
7 a 2 c 7 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred.No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 12 GCCTCT 7

RESULT 10
A2800646 19 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0058H14R Mouse 10kb plasmid UGC1M library Mus musculus genomic
DEFINITION clone UGC2M0058H14 R, DNA sequence.
ACCESSION A2800646
VERSION A2800646.1 GI:12952969
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0058 row: H column: 14
Seq primer: CACACAGGAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 19.
location/Qualifiers

FEATURES
source

1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U062M0058H14"
/clone_1lb="Mouse 10kb plasmid U062M0058H14"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 4 c 7 g 5 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GCCTCTC 6
|||||
Db 7 GCCTCT 2

RESULT 11
A2834038 19 bp DNA linear GSS 20-FEB-2001
LOCUS A2834038.1
DEFINITION clone U062M0116H01 R, DNA sequence.
ACCESSION A2834038
VERSION A2834038.1 GI:13003946
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0116 row: H column: 01
Seq primer: CACACAGGAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 19.
location/Qualifiers

FEATURES
source

1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U062M0116H01"
/clone_1lb="Mouse 10kb plasmid U062M0116H01"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 2 c 7 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 8.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GCCTCTC 6
|||||
Db 6 GCCTCT 1

RESULT 12
BQ789776 20 bp mRNA linear EST 30-JUL-2002
LOCUS BQ789776
DEFINITION hage002a10 Heterobasidion annosum - Scots pine infection stage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage002a10, mRNA sequence.
ACCESSION BQ789776
VERSION BQ789776.1 GI:22004738
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 20)

AUTHORS Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
TITLE Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (*Heterobasidion annosum*) with seedling roots of Scots pine (*Pinus sylvestris*)
JOURNAL Unpublished (2001)
COMMENT Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden
Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer.
FEATURES
source Location/Qualifiers
1.20
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone="hage002a10"
/clone_lib="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with *H. annosum*"
/note="Vector: pT-Adv; Site_1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of *Heterobasidion annosum* (Fp5)."
BASE COUNT 4 a 5 c 3 g 7 t 1 others
ORIGIN
Query Match 100.0%; Score 6; DB 14; Length 20;
Best local Similarity 83.3%; Pred. No. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCT 6
|||||
DB 12 GCCTCT 17
RESULT 13
AZ309156 20 bp DNA linear GSS 29-SEP-2000
LOCUS IM0013B09F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
DEFINITION clone UUGCIM0013B09 F, DNA sequence.
ACCESSION AZ309156
VERSION AZ309156.1 GI:10349862
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weis, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0013 row: B column: 09
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 20.
FEATURES
source Location/Qualifiers
1.20

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCIM0013B09"
/clone_lib="Mouse 10kb plasmid UUGCIM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD2env; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321149b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 2 a 11 c 4 g 3 t
ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 20;
Best local Similarity 83.3%; Pred. No. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCT 6
|||||
DB 3 GCCTCT 8
RESULT 14
AZ366535 20 bp DNA linear GSS 02-OCT-2000
LOCUS IM0115M15R Mouse 10kb plasmid UUGCIM library Mus musculus genomic
DEFINITION clone UUGCIM0115M15 R, DNA sequence.
ACCESSION AZ366535
VERSION AZ366535.1 GI:10480235
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weis, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0115 row: M column: 15
Seq primer: CACACAGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
FEATURES
source Location/Qualifiers
1.20

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1147321141gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      5 a      6 c      3 g      6 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
|||||
Db 14 GCCTCT 19

```

```

RESULT 15
LOCUS      A2615164      20 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION 1M0444B24F Mouse 10kb plasmid UGCGM library Mus musculus genomic
clone UGCGM0444B24 F, DNA sequence.
ACCESSION  A2615164
VERSION    A2615164
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE  1 (bases 1 to 20)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177

```

```

FEATURES
Source     Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0444 row: B column: 24
            Seq primer: CGTGTGTAACGACGCGCCAGT
            Class: plasmid
            High quality sequence stop: 20.
            Location/Qualifiers
            1..20

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1147321141gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      3 a      7 c      5 g      5 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 8.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
|||||
Db 7 GCCTCT 12

```

```

Search completed: May 21, 2003, 08:37:46
Job time : 1093.5 secs

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 03:56:27 ; Search time 333.5 Seconds

(Without alignments)
523.589 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6
Sequence: 1 gcctcu 6

Scoring table: IDENTITY_NDC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_or:*

21: em_ov:*

22: em_ov:*

23: em_pal:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pin:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_higo_hum:*

40: em_higo_mus:*

41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| C 1 | 6 | 100.0 | 10 | 6 | ARI28987 | ARI28987 Sequence |
| C 2 | 6 | 100.0 | 10 | 6 | ARI28988 | ARI28988 Sequence |
| C 3 | 6 | 100.0 | 10 | 6 | ARI28989 | ARI28989 Sequence |
| C 4 | 6 | 100.0 | 10 | 6 | ARI28990 | ARI28990 Sequence |
| C 5 | 6 | 100.0 | 10 | 6 | ARI28991 | ARI28991 Sequence |
| C 6 | 6 | 100.0 | 10 | 6 | AX152349 | AX152349 Sequence |
| C 7 | 6 | 100.0 | 10 | 6 | AX152410 | AX152410 Sequence |
| C 8 | 6 | 100.0 | 10 | 6 | AX152849 | AX152849 Sequence |
| C 9 | 6 | 100.0 | 10 | 6 | AX152886 | AX152886 Sequence |
| C 10 | 6 | 100.0 | 10 | 6 | AX153384 | AX153384 Sequence |
| C 11 | 6 | 100.0 | 10 | 6 | AX153513 | AX153513 Sequence |
| C 12 | 6 | 100.0 | 10 | 6 | E16890 | E16890 DNA sequenc |
| C 13 | 6 | 100.0 | 11 | 6 | A91501 | A91501 Sequence 28 |
| C 14 | 6 | 100.0 | 11 | 6 | A91503 | A91503 Sequence 30 |
| C 15 | 6 | 100.0 | 11 | 6 | AX470495 | AX470495 Sequence |
| C 16 | 6 | 100.0 | 11 | 6 | AX470740 | AX470740 Sequence |
| C 17 | 6 | 100.0 | 11 | 6 | AX470747 | AX470747 Sequence |
| C 18 | 6 | 100.0 | 11 | 6 | AX471502 | AX471502 Sequence |
| C 19 | 6 | 100.0 | 11 | 6 | AX471630 | AX471630 Sequence |
| C 20 | 6 | 100.0 | 11 | 6 | AX471677 | AX471677 Sequence |
| C 21 | 6 | 100.0 | 11 | 6 | AX471805 | AX471805 Sequence |
| C 22 | 6 | 100.0 | 12 | 6 | A06058 | A06058 Synthetic p |
| C 23 | 6 | 100.0 | 12 | 6 | A06059 | A06059 Synthetic p |
| C 24 | 6 | 100.0 | 12 | 6 | A15123 | A15123 Nucleotide |
| C 25 | 6 | 100.0 | 12 | 6 | A16601 | A16601 Nucleotide |
| C 26 | 6 | 100.0 | 12 | 6 | A16602 | A16602 Nucleotide |
| C 27 | 6 | 100.0 | 12 | 6 | A47656 | A47656 Sequence 16 |
| C 28 | 6 | 100.0 | 12 | 6 | A61520 | A61520 Sequence 89 |
| C 29 | 6 | 100.0 | 12 | 6 | A91489 | A91489 Sequence 16 |
| C 30 | 6 | 100.0 | 12 | 6 | AR027874 | AR027874 Sequence |
| C 31 | 6 | 100.0 | 12 | 6 | AR075465 | AR075465 Sequence |
| C 32 | 6 | 100.0 | 12 | 6 | AR075468 | AR075468 Sequence |
| C 33 | 6 | 100.0 | 12 | 6 | ARI01000 | ARI01000 Sequence |
| C 34 | 6 | 100.0 | 12 | 6 | ARI37925 | ARI37925 Sequence |
| C 35 | 6 | 100.0 | 12 | 6 | ARI53916 | ARI53916 Sequence |
| C 36 | 6 | 100.0 | 12 | 6 | ARI53919 | ARI53919 Sequence |
| C 37 | 6 | 100.0 | 12 | 6 | ARI67701 | ARI67701 Sequence |
| C 38 | 6 | 100.0 | 12 | 6 | ARI78533 | ARI78533 Sequence |
| C 39 | 6 | 100.0 | 12 | 6 | ARI78536 | ARI78536 Sequence |
| C 40 | 6 | 100.0 | 12 | 6 | ARI99100 | ARI99100 Sequence |
| C 41 | 6 | 100.0 | 12 | 6 | AX233626 | AX233626 Sequence |
| C 42 | 6 | 100.0 | 12 | 6 | E29585 | E29585 Method for |
| C 43 | 6 | 100.0 | 12 | 6 | E38691 | E38691 Method and |
| C 44 | 6 | 100.0 | 12 | 6 | E64117 | E64117 Method for |
| C 45 | 6 | 100.0 | 12 | 6 | I07918 | I07918 Sequence 30 |

ALIGNMENTS

RESULT 1

LOCUS ARI28987/c

DEFINITION Sequence 2 from patent US 6183966.

ACCESSION ARI28987

VERSION ARI28987.1 GI:14116649

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unknown.

REFERENCE 1 (bases 1 to 10)

AUTHORS Gray,D.M. and Clark,C.L.

TITLE Apparatus and method for selectively ranking sequences for antisense targeting

JOURNAL Patent: US 6183966-A 2 06-FEB-2001;

Pred. No. is the number of results predicted by chance to have a

FEATURES Location/Qualifiers
source 1..10 /organism="unknown"
BASE COUNT 2 a 3 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCTCU 6
DB 6 GCCCTCT 1

RESULT 2
LOCUS AR128988/c 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6183966.
ACCESSION AR128988
VERSION AR128988.1 GI:14116650
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURN. Patent: US 6183966-A 3 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10 /organism="unknown"

BASE COUNT 2 a 3 c 5 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCTCU 6
DB 7 GCCCTCT 2

RESULT 3
LOCUS AR128989/c 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 4 from patent US 6183966.
ACCESSION AR128989
VERSION AR128989.1 GI:14116651
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURN. Patent: US 6183966-A 4 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10 /organism="unknown"

BASE COUNT 3 a 3 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCTCU 6
DB 8 GCCCTCT 3

RESULT 4
LOCUS AR128990/c 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6183966.
ACCESSION AR128990
VERSION AR128990.1 GI:14116652
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURN. Patent: US 6183966-A 5 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10 /organism="unknown"

BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCTCU 6
DB 9 GCCCTCT 4

RESULT 5
LOCUS AR128991/c 10 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6183966.
ACCESSION AR128991
VERSION AR128991.1 GI:14116653
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Gray,D.M. and Clark,C.L.
TITLE Apparatus and method for selectively ranking sequences for antisense targeting
JOURN. Patent: US 6183966-A 6 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..10 /organism="unknown"

BASE COUNT 4 a 2 c 4 g 0 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCTCU 6
DB 10 GCCCTCT 5

RESULT 6
LOCUS AX152349 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 264 from patent W00138577.
ACCESSION AX152349
VERSION AX152349.1 GI:14534000
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 264 31-MAY-2001;
FEATURES
    source
        1.10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 0 a 5 c 2 g 3 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
Db 1 GCCTCT 6

RESULT 7
AX152410 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS Sequence 325 from Patent WO0138577.
DEFINITION AX152410
ACCESSION AX152410
VERSION AX152410.1 GI:14534061
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 325 31-MAY-2001;
FEATURES
    source
        1.10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 1 a 5 c 2 g 2 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
Db 1 GCCTCT 6

RESULT 8
AX152849 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS Sequence 764 from Patent WO0138577.
DEFINITION AX152849
ACCESSION AX152849
VERSION AX152849.1 GI:14534500
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 764 31-MAY-2001;
FEATURES
    source
        1.10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
Db 1 GCCTCT 6

RESULT 9
AX152886/c 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS Sequence 801 from Patent WO0138577.
DEFINITION AX152886
ACCESSION AX152886
VERSION AX152886.1 GI:14534537
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 801 31-MAY-2001;
FEATURES
    source
        1.10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 5 a 1 c 3 g 1 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
Db 8 GCCTCT 3

RESULT 10
AX153384 10 bp DNA 11linear PAT 22-JUN-2001
LOCUS Sequence 1299 from Patent WO0138577.
DEFINITION AX153384
ACCESSION AX153384
VERSION AX153384.1 GI:14535035
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 1299 31-MAY-2001;
FEATURES
    source
        1.10
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 0 a 4 c 3 g 3 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
Db 1 GCCTCT 6
```

Db 2 GCCTCT 7

RESULT 11
AX153513
LOCUS AX153513 10 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1428 from Patent WO0138577.
ACCESSION AX153513
VERSION AX153513.1 GI:14535164
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 10)
TITLE Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
JOURNAL Human transcriptomes
Patent: WO 0138577-A 1428 31-MAY-2001;
The Johns Hopkins University (US)
Location/Qualifiers
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 0 a 4 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 1 GCCTCT 6

RESULT 12
E16890 10 bp DNA linear PAT 28-JUL-1999
LOCUS E16890
DEFINITION DNA sequence required for efficient protein transcription in
Brevibacterium flavum.
ACCESSION E16890
VERSION E16890.1 GI:5711573
KEYWORDS JP 1998229881-A/31.
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS 1 (bases 1 to 10)
TITLE Kobayashi,M., Men,T. and Yugawa,H.
JOURNAL DNA HAVING SEQUENCE CAPABLE OF EFFICIENTLY TRANSLATING PROTEIN IN
CORYNEFORM BACTERIA
Patent: JP 1998229881-A 31 02-SEP-1998;
MITSUBISHI CHEM CORP
COMMENT
OS Brevibacterium flavum
PN JP 1998229881-A/31
PD 02-SEP-1998
PF 19-FEB-1997 JP 1997035338
PI KOBAYASHI MIKI, MAN TOMOKO, YUGAWA HIDEAKI
PC C12N15/09,C07H21/04,C12N1/21//C12N9/38,C12Q1/68,(C12N15/09, PC
C12N1:19),
PC (C12N1/21,C12N1:13),(C12N9/38,C12N1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH key
FT source 1..10
/organism="Brevibacterium flavum" FT
/strain="MJ-233"
Location/Qualifiers
FEATURES
source 1..10
/organism="Corynebacterium glutamicum"

BASE COUNT 3 a 1 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 9 GCCTCT 4

RESULT 13
A91501 11 bp DNA linear PAT 22-JAN-2000
LOCUS A91501
DEFINITION Sequence 28 from Patent WO9824928.
ACCESSION A91501
VERSION A91501.1 GI:6740456
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
AUTHORS 1 (bases 1 to 11)
TITLE Pallsgaard,N. and Hokland,P.
JOURNAL DETECTION OF CHROMOSOMAL ABNORMALITIES
Patent: WO 9824928-A 28 11-JUN-1998;
PALLSGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
FEATURES
source 1..11
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1 a 5 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 5 GCCTCT 10

RESULT 14
A91503 11 bp DNA linear PAT 22-JAN-2000
LOCUS A91503
DEFINITION Sequence 30 from Patent WO9824928.
ACCESSION A91503
VERSION A91503.1 GI:6740458
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
AUTHORS 1 (bases 1 to 11)
TITLE Pallsgaard,N. and Hokland,P.
JOURNAL DETECTION OF CHROMOSOMAL ABNORMALITIES
Patent: WO 9824928-A 30 11-JUN-1998;
PALLSGAARD NIELS (DK); HOKLAND PETER (DK)
Location/Qualifiers
FEATURES
source 1..11
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 0 a 5 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. No. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 5 GCCTCT 10

```

RESULT 15
AX470495/c  AX470495  11 bp  DNA  linear  PAT 09-AUG-2002
LOCUS      Sequence 72 from Patent WO02053773.
ACCESSION  AX470495
VERSION    AX470495.1  GI:22205620
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS   Hofmann,K., Conradt,M. and Petersohn,D.
TITLE     Method for determining skin stress or skin ageing in vitro
JOURNAL   Patent: WO 02053773-A 72 11-JUL-2002;
            HENKEL KGAA (DE)
FEATURES
            Location/Qualifiers
            source          1..11
                               /organism="Homo sapiens"
                               /db_xref="taxon:9606"
BASE COUNT  6 a          2 c          3 g          0 t
ORIGIN
Query Match          100.0%; Score 6; DB 6; Length 11;
Best Local Similarity 83.3%; Pred. NO. 2e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY      1 GCCTCU 6
        |||||:
Db      8 GCCTCT 3

Search completed: May 21, 2003, 05:12:46
Job time : 334 secs

```


PT Subset of primers able to amplify group of related sequences
 XX
 PS Claim 21; Page 75; 92pp; English.
 XX
 CC The invention provides primers (AAx29501-x29679) for identifying
 CC sequences encoding structurally or functionally related proteins such as
 CC nuclear or G-protein coupled receptors, apoptosis-related or DNA
 CC repair/replication proteins. The identified sequences are broadly useful
 CC in plant biology, agriculture, human or veterinary medicine,
 CC reproduction, microbiology or environmental science, e.g. to study
 CC expression of nuclear receptors at different stages of tissue development
 CC or after treatment with particular drugs. It is also used for DNA
 CC fingerprinting (to generate products useful for differential
 CC hybridisation), or, where a 3'-anchor primer is used, to isolate the
 CC 3'-ends of mRNA sequences. Sequences AAx29526-x29583 represent claimed
 CC primers specific for human G-protein coupled receptor genes.
 XX
 SQ Sequence 8 BP; 0 A; 4 C; 2 G; 2 T; 0 other:
 Query Match 100.0%; Score 6; DB 20; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.7e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DY 1 GCCTCU 6
 DB 1 GCCTCT 6
 RESULT 2
 ID AAA80736
 AC AAA80736 standard; DNA: 8 BP.
 XX
 AC AAA80736;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE A. thaliana primer walking octamer SEQ ID NO: 49.
 XX
 KW Primer walking; octamer; primer; DNA sequencing; PCR; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US6083695-A.
 XX
 PD 04-JUL-2000.
 XX
 PF 21-MAY-1997; 97US-0859954.
 XX
 PR 15-APR-1996; 96US-0632782.
 XX
 PA (UYHO-) UNIV HOUSTON.
 PA (HARD/) HARDIN S H.
 XX
 PI Hardin PE, Hardin SH, Homayouni R;
 DR WPI: 2000-474852/41.
 XX
 XX Sequencing an unknown DNA molecule for the polymerase chain reaction
 PT and other primer processes comprises primer walking of octamer
 PT oligonucleotides -
 XX
 XX Example 8; Column 49-50; 16pp; English.
 PS
 XX
 CC This invention describes a novel method for sequencing an unknown DNA
 CC molecule which comprises selecting a library primer from an octamer
 CC oligonucleotide library consisting of 48 8-bp sequences and
 CC corresponding complementary sequences, where the library primer is
 CC complementary to a known sequence adjacent to the unknown sequence or
 CC is complementary to a sequence in a known extension product. The method
 CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
 CC which make use of primers. The octamers are used to identify coding
 CC sequences. Primer walking using the octamer libraries is advantageous
 CC over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
 CC directed and methodical approach. AAA80688-A81253 represent the octamer
 CC primers used in the primer walking method of the invention.
 XX
 SQ Sequence 8 BP; 0 A; 3 C; 1 G; 4 T; 0 other:
 Query Match 100.0%; Score 6; DB 21; Length 8;
 Best Local Similarity 83.3%; Pred. No. 2.7e+08;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DY 1 GCCTCU 6
 DB 3 GCCTCT 8
 RESULT 3
 ID ABQ71469/C
 AC ABQ71469 standard; DNA: 9 BP.
 XX
 AC ABQ71469;
 XX
 DT 28-AUG-2002 (first entry)
 XX
 DE zinc finger protein related oligonucleotide target SEQ ID NO:588.
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN W0200242459-A2.
 XX
 PD 30-MAY-2002.
 XX
 PE 20-NOV-2001; 2001MO-US43438.
 XX
 PR 20-NOV-2000; 2000US-0716637.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 PA
 PI Liu Q;
 DR WPI: 2002-500284/53.
 XX
 XX New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering.
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 XX
 PS Example 1; Page 44; 81pp; English.
 XX
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (I) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 CC that it binds to the S2 target subsite, and selecting the F3 zinc
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABQ48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

XX Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other;
SQ

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 4
ABQ71786/c
ID ABQ71786 standard; DNA; 9 BP.

XX ABQ71786;

XX 28-AUG-2002 (first entry)

DE zinc finger protein related oligonucleotide target SEQ ID NO:2084.

XX zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

XX Homo sapiens.

OS Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -

XX Example 1; Page 55; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 24; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.4e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 5
ABQ71788/c
ID ABQ71788 standard; DNA; 9 BP.

XX ABQ71788;

XX 28-AUG-2002 (first entry)

DE zinc finger protein related oligonucleotide target SEQ ID NO:2086.

XX zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

XX Homo sapiens.

OS Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus -

XX Example 1; Page 55; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
CC that it binds to the S2 target sub-site, and selecting the F3 zinc
CC finger such that it binds to the S3 target sub-site, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target sub-sites having the nucleotide G in the 5'-most position of the
CC sub-site. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX Sequence 9 BP; 3 A; 1 C; 4 G; 1 T; 0 other;

XX Query Match 100.0%; Score 6; DB 24; Length 9;
XX Best Local Similarity 83.3%; Pred. No. 2.4e+08;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 6

ABQ71921/c
ID ABQ71921 standard; DNA; 9 BP.

AC ABQ71921;

DE 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2219.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN MO200242459-A2.

PD 30-MAY-2002.

PF 20-NOV-2001; 2001MO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying

PT gene function and for human therapeutics and plant engineering,

PT comprises first, second and third zinc fingers, ordered from N- to

PT C-terminus

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Db 8 GCCTCT 3

RESULT 7

ABQ71982/c
ID ABQ71982 standard; DNA; 9 BP.

AC ABQ71982;

DE 28-AUG-2002 (first entry)

DE Zinc finger protein related oligonucleotide target SEQ ID NO:2280.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.

OS Homo sapiens.

OS Synthetic.

PN MO200242459-A2.

PD 30-MAY-2002.

PF 20-NOV-2001; 2001MO-US43438.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Liu Q;

DR WPI; 2002-500284/53.

PT New zinc finger protein that binds to target site, useful in studying

PT gene function and for human therapeutics and plant engineering,

PT comprises first, second and third zinc fingers, ordered from N- to

PT C-terminus

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 9 BP; 4 A; 1 C; 4 G; 0 U; 0 other:

Query Match 100.0%; Score 6; DB 24; Length 9;

Best Local Similarity 83.3%; Pred. No. 2.4e+08; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0;

Db 1 GCCTCU 6

|||||:
8 GCCTCT 3

```

RESULT 8
ABO72173/C
ID ABO72173 standard; DNA; 9 BP.
XX
XX ABO72173;
AC
XX
XX 28-AUG-2002 (first entry)
DT
XX
XX Zinc finger protein related oligonucleotide target SEQ ID NO:2471.
DE
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX
XX Homo sapiens.
OS
XX
XX Synthetic.
PN
XX WO200242459-A2.
XX
XX 30-MAY-2002.
PD
XX
XX 20-NOV-2001; 2001WO-US43438.
PF
XX
XX 20-NOV-2000; 2000US-0716637.
PR
XX
XX 20-NOV-2000; 2000US-0716637.
PA
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
PI
XX
XX Liu Q;
DR
XX
XX WPI; 2002-500284/53.
XX
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus
XX
XX
XX Example 1; Page 63; 81pp; English.
XX
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (1); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABO71213 to ABO72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
XX Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
SQ
XX
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 6; DB 24; Length 9;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX OY 1 GCCTCU 6
XX |||||:
XX Db 8 GCCTCT 3
XX
XX
XX RESULT 9
XX ABO72175/C
XX ID ABO72175 standard; DNA; 9 BP.
XX

```

```

XX
XX ABO72175;
AC
XX
XX 28-AUG-2002 (first entry)
DT
XX
XX Zinc finger protein related oligonucleotide target SEQ ID NO:2473.
DE
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger; ss.
KM
XX
XX Homo sapiens.
OS
XX
XX Synthetic.
PN
XX WO200242459-A2.
XX
XX 30-MAY-2002.
PD
XX
XX 20-NOV-2001; 2001WO-US43438.
PF
XX
XX 20-NOV-2000; 2000US-0716637.
PR
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA
XX
XX Liu Q;
DR
XX
XX WPI; 2002-500284/53.
XX
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus
XX
XX
XX Example 1; Page 63; 81pp; English.
XX
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (1); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determined the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABO71213 to ABO72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX
XX Sequence 9 BP; 2 A; 1 C; 5 G; 1 T; 0 other;
SQ
XX
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 6; DB 24; Length 9;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX OY 1 GCCTCU 6
XX |||||:
XX Db 8 GCCTCT 3
XX
XX
XX RESULT 10
XX AAQ96863/C
XX ID AAQ96863 standard; DNA; 10 BP.
XX

```

DT 26-MAR-1996 (first entry)
 XX
 DE HIV-1 NLA-3 nef gene nucleotide deletion 458.
 XX
 XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PF 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.
 XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX WPI; 1995-293115/38.
 DR
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 PS Claim 13; Page 194; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NLA-3 involves deletion of 1
 CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
 CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 1 C; 4 G; 1 T; 0 other;
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 83.3%; Pred. No. 2.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCU 6
 DB 10 GCCTCT 5
 RESULT 11
 AA096864/C
 ID AA096864 standard; DNA: 10 BP.
 XX
 AC AA096864;
 XX
 DT 26-MAR-1996 (first entry)
 XX
 DE HIV-1 NLA-3 nef gene nucleotide deletion 459.
 XX
 KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PF 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.

XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
 XX
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX WPI; 1995-293115/38.
 DR
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 PS Claim 13; Page 194; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NLA-3 involves deletion of 1
 CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
 CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region; the
 CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
 CC (AA096141). The resulting avirulent HIV strains are still capable of
 CC inducing an immune response in humans, and enable the generation of
 CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
 XX
 SQ Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;
 Query Match 100.0%; Score 6; DB 16; Length 10;
 Best Local Similarity 83.3%; Pred. No. 2.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTCU 6
 DB 9 GCCTCT 4
 RESULT 12
 AA096865/C
 ID AA096865 standard; DNA: 10 BP.
 XX
 AC AA096865;
 XX
 DT 26-MAR-1996 (first entry)
 XX
 DE HIV-1 NLA-3 nef gene nucleotide deletion 460.
 XX
 KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
 OS Human immunodeficiency virus type 1.
 XX
 PN WO9521912-A1.
 XX
 PD 17-AUG-1995.
 XX
 PF 14-FEB-1995; 95WO-AU00063.
 XX
 PR 23-DEC-1994; 94AU-0000284.
 PR 14-FEB-1994; 94AU-0003864.
 PR 21-FEB-1994; 94AU-0004002.
 XX
 PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
 PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
 XX WPI; 1995-293115/38.
 DR
 XX
 PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
 PT or LTR region - can be used in a vaccine to inhibit/reduce
 PT productive infection in an individual by a pathogenic strain
 PS Claim 13; Page 194; 301pp; English.
 XX
 CC Attenuation of pathogenic HIV-1 strain NLA-3 involves deletion of 1
 CC or more decanucleotides (AA096406-Q97018) from the nef gene and/or
 CC 1 or more decanucleotides (AA097019-Q97166) from the LTR region; the

CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene (AAQ96141). The resulting avirulent HIV strains are still capable of inducing an immune response in humans, and enable the generation of CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

SQ Sequence 10 BP; 4 A; 2 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 83.3%; Pred. No. 2.7e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6

DB 8 GCCTCT 3

RESULT 13

AAQ96866/c

ID AAQ96866 standard; DNA: 10 BP.

XX AAQ96866;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 461.

KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

PD 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PA (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

DR WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

PS Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deannucleotides (AAQ96406-097018) from the nef gene and/or

CC 1 or more deannucleotides (AAQ97019-097166) from the LTR region; the

CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene

CC (AAQ96141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

SQ Sequence 10 BP; 5 A; 2 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 83.3%; Pred. No. 2.7e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6

DB 7 GCCTCT 2

RESULT 14

AAQ96867/c

ID AAQ96867 standard; DNA: 10 BP.

XX AAQ96867;

DT 26-MAR-1996 (first entry)

DE HIV-1 NL4-3 nef gene nucleotide deletion 462.

KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

OS Human immunodeficiency virus type 1.

PN WO9521912-A1.

PD 17-AUG-1995.

PF 14-FEB-1995; 95WO-AU00063.

PR 23-DEC-1994; 94AU-0000284.

PR 14-FEB-1994; 94AU-0003864.

PR 21-FEB-1994; 94AU-0004002.

PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

PA (MACF-) MACFARLANE BURNET CENT MEDICAL.

PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

DR WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

PT or LTR region - can be used in a vaccine to inhibit/reduce

PT productive infection in an individual by a pathogenic strain

PS Claim 13; Page 194; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

CC or more deannucleotides (AAQ96406-097018) from the nef gene and/or

CC 1 or more deannucleotides (AAQ97019-097166) from the LTR region; the

CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene

CC (AAQ96141). The resulting avirulent HIV strains are still capable of

CC inducing an immune response in humans, and enable the generation of

CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

SQ Sequence 10 BP; 4 A; 2 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 6; DB 16; Length 10;

Best Local Similarity 83.3%; Pred. No. 2.7e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6

DB 6 GCCTCT 1

RESULT 15

AAQ96201

ID AAQ96201 standard; DNA: 10 BP.

XX AAQ96201;

DT 22-SEP-1999 (first entry)

DE SAGE tag used to identify transcripts which are enhanced by p53.

XX p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity;

KW neoplastic; p53 binding site; p53 promoter; SAGE tag; ss.

XX Homo sapiens.

OS WO9914356-A2.

PN 25-MAR-1999.

PD

```

XX 17-SEP-1998; 98WO-US19300.
PF
XX
PR 30-MAR-1998; 98US-0079817.
PR 17-SEP-1997; 97US-0059153.
XX
PA (UWJO ) UNIV JOHNS HOPKINS.
XX
PI
XX
PI
XX
DR WPI: 1999-443793/37.
XX
PT
XX
PT
XX
PS Use of p53 transcription tags to determine p53 status in, e.g.
    cancer diagnosis
    Example 1; Page 25; 73pp; English.
XX
CC The specification describes the use of p53 transcription tags for
CC developing products to determine p53 status, to diagnose cancer
CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
CC A method for diagnosing cancer or determining p53 status in a sample
CC suspected for being neoplastic comprises comparing the level of
CC transcription of an RNA transcript in a first sample (s1) of a first
CC tissue (t1) to the level of transcription of the transcript in a second
CC sample (s2) of a second tissue (s2), where s1 is suspected of being
CC neoplastic and s2 is a normal human tissue (of the same type) and the
CC transcript is identified by a tag; and categorizing s1 as neoplastic
CC or as having a mutant p53 when transcription is found to be the same
CC or lower in the first, than in s2. The methods and products can be used
CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
CC or carcinogenicity of a test agent. AXX86201-33 represent SAGE tags
CC used to identify transcripts which are enhanced by p53.
XX
SQ Sequence 10 BP; 0 A; 6 C; 1 G; 3 T; 0 other;
    Query Match 100.0%; Score 6; DB 20; Length 10;
    Best Local Similarity 83.3%; Pred. No. 2.7e+05;
    Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
    |||||
Db 4 GCCTCT 9

```

Search completed: May 21, 2003, 04:50:11.
 Job time : 156 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:37:42 ; Search time 35.25 Seconds
(without alignments)
52.200 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6
Sequence: 1 gcctcu 6

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCBUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfileseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1 | 6 | 100.0 | 8 | US-08-859-954-49 | Sequence 49, App1 |
| 2 | 6 | 100.0 | 10 | US-08-388-353-459 | Sequence 49, App |
| 3 | 6 | 100.0 | 10 | US-08-388-353-460 | Sequence 460, App |
| 4 | 6 | 100.0 | 10 | US-08-388-353-461 | Sequence 461, App |
| 5 | 6 | 100.0 | 10 | US-08-388-353-462 | Sequence 462, App |
| 6 | 6 | 100.0 | 10 | US-08-388-353-463 | Sequence 463, App |
| 7 | 6 | 100.0 | 10 | US-08-488-551B-459 | Sequence 459, App |
| 8 | 6 | 100.0 | 10 | US-08-488-551B-460 | Sequence 460, App |
| 9 | 6 | 100.0 | 10 | US-08-488-551B-461 | Sequence 461, App |
| 10 | 6 | 100.0 | 10 | US-08-488-551B-462 | Sequence 462, App |
| 11 | 6 | 100.0 | 10 | US-08-488-551B-463 | Sequence 463, App |
| 12 | 6 | 100.0 | 10 | US-09-235-614-2 | Sequence 2, App1 |
| 13 | 6 | 100.0 | 10 | US-09-235-614-3 | Sequence 3, App1 |
| 14 | 6 | 100.0 | 10 | US-09-235-614-4 | Sequence 4, App1 |
| 15 | 6 | 100.0 | 10 | US-09-235-614-5 | Sequence 5, App1 |
| 16 | 6 | 100.0 | 10 | US-09-235-614-6 | Sequence 6, App1 |
| 17 | 6 | 100.0 | 10 | US-09-154-750A-4 | Sequence 4, App1 |
| 18 | 6 | 100.0 | 10 | US-08-494-301A-16 | Sequence 16, App1 |
| 19 | 6 | 100.0 | 12 | US-08-480-020B-18 | Sequence 18, App1 |
| 20 | 6 | 100.0 | 12 | US-08-480-020B-21 | Sequence 21, App1 |
| 21 | 6 | 100.0 | 12 | US-08-910-618-18 | Sequence 18, App1 |
| 22 | 6 | 100.0 | 12 | US-08-910-618-21 | Sequence 21, App1 |
| 23 | 6 | 100.0 | 12 | US-08-779-335-5 | Sequence 5, App1 |
| 24 | 6 | 100.0 | 12 | US-08-671-824-19 | Sequence 19, App1 |
| 25 | 6 | 100.0 | 12 | US-08-874-825-88 | Sequence 88, App1 |
| 26 | 6 | 100.0 | 12 | US-08-938-835A-5 | Sequence 5, App1 |
| 27 | 6 | 100.0 | 12 | US-08-663-824-88 | Sequence 88, App1 |

| | | | | | |
|----|---|-------|----|--------------------|-------------------|
| 28 | 6 | 100.0 | 12 | US-09-243-335-1 | Sequence 1, App1 |
| 29 | 6 | 100.0 | 12 | US-08-910-322-18 | Sequence 18, App1 |
| 30 | 6 | 100.0 | 12 | US-08-910-322-21 | Sequence 21, App1 |
| 31 | 6 | 100.0 | 12 | US-09-281-418-65 | Sequence 65, App1 |
| 32 | 6 | 100.0 | 12 | US-08-484-939A-18 | Sequence 18, App1 |
| 33 | 6 | 100.0 | 12 | US-08-484-939A-21 | Sequence 21, App1 |
| 34 | 6 | 100.0 | 12 | US-09-043-149-48 | Sequence 48, App1 |
| 35 | 6 | 100.0 | 12 | US-09-231-303-88 | Sequence 88, App1 |
| 36 | 6 | 100.0 | 13 | US-08-284-746-5 | Sequence 5, App1 |
| 37 | 6 | 100.0 | 13 | US-08-284-746-16 | Sequence 16, App1 |
| 38 | 6 | 100.0 | 13 | US-08-050-073-162 | Sequence 162, App |
| 39 | 6 | 100.0 | 14 | US-08-303-004-23 | Sequence 23, App1 |
| 40 | 6 | 100.0 | 14 | US-08-050-073-99 | Sequence 99, App1 |
| 41 | 6 | 100.0 | 14 | US-08-050-073-161 | Sequence 161, App |
| 42 | 6 | 100.0 | 14 | US-08-192-941-25 | Sequence 25, App |
| 43 | 6 | 100.0 | 14 | US-08-171-718-119 | Sequence 119, App |
| 44 | 6 | 100.0 | 14 | US-08-173-489C-324 | Sequence 324, App |
| 45 | 6 | 100.0 | 2 | US-08-232-087A-6 | Sequence 6, App1 |

ALIGNMENTS

RESULT 1
US-08-859-954-49
Sequence 49, Application US/08859954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramin
TITLE OF INVENTION: Design and Optimized Primer Library for
NUMBER OF INVENTION: Gene Sequencing and Method Thereof
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,954
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D. 32,714
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
HYPOTHETICAL: YES
ANTI-SENSE: YES
US-08-859-954-49
Query Match 100.0%; Score 6; DB 3; Length 8;

Best Local Similarity 83.3%; Pred. NO. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCTCU 6
|||||:
Db 3 GCCTCT 8

RESULT 2
US-08-388-353-459/C

Sequence 459, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 10 GCCTCT 5

RESULT 3

US-08-388-353-460/C
Sequence 460, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. NO. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||:
Db 9 GCCTCT 4

RESULT 4

US-08-388-353-461/C
Sequence 461, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 5
US-08-388-353-462/C
Sequence 462, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-462

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 7 GCCTCT 2

RESULT 6
US-08-388-353-463/C
Sequence 463, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-463

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
|||||
DB 6 GCCTCT 1

RESULT 7
US-08-488-551B-459/C
Sequence 459, Application US/0848851B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 459:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-459

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
DB 10 GCCCTC 5

RESULT 8
US-08-488-551B-460/C
Sequence 460, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 460:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-460

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTC 6
DB 9 GCCCTC 4

RESULT 9
US-08-488-551B-461/C
Sequence 461, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 461:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-461

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

QY 1 GCCTCU 6
DB 8 GCCTCT 3

RESULT 10
US-08-488-551B-462/C
Sequence 462, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-462
Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
DB 7 GCCTCT 2

RESULT 11
US-08-488-551B-463/C
Sequence 463, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 463:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-463
Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTCU 6
DB 6 GCCTCT 1

RESULT 12
US-09-235-614-2/C
Sequence 2, Application US/09235614
Patent No. 6183966
GENERAL INFORMATION:
APPLICANT: GRAY, DONALD M.

```

: APPLICANT: CLARK, CHRISTOPHER L.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 2
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-2

```

```

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 GCCTCU 6
DB 6 GCCTCT 1

```

```

RESULT 13
US-09-235-614-3/c
: Sequence 3, Application US/09235614
: Patent No. 6183966
: GENERAL INFORMATION:
: APPLICANT: GRAY, DONALD M.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 3
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-3

```

```

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 GCCTCU 6
DB 7 GCCTCT 2

```

```

RESULT 14
US-09-235-614-4/c
: Sequence 4, Application US/09235614
: Patent No. 6183966
: GENERAL INFORMATION:
: APPLICANT: GRAY, DONALD M.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 4
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

```

```

: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 4
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-4

```

```

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 GCCTCU 6
DB 8 GCCTCT 3

```

```

RESULT 15
US-09-235-614-5/c
: Sequence 5, Application US/09235614
: Patent No. 6183966
: GENERAL INFORMATION:
: APPLICANT: GRAY, DONALD M.
: TITLE OF INVENTION: AN APPARATUS AND METHOD FOR SELECTIVELY RANKING
: FILE REFERENCE: 91556/66384
: CURRENT APPLICATION NUMBER: US/09/235,614
: CURRENT FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 08/808,474
: PRIOR FILING DATE: 1997-03-03
: PRIOR APPLICATION NUMBER: 08/320,507
: PRIOR FILING DATE: 1994-10-07
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 5
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: ASO
US-09-235-614-5

```

```

Query Match          100.0%; Score 6; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 GCCTCU 6
DB 9 GCCTCT 4

```

```

Search completed: May 21, 2003, 06:28:35
Job time : 35.25 secs

```

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:13:38 ; Search time 1094 Seconds
(without alignments)
88.824 Million cell updates/sec

Title: US-09-936-146-9

Perfect score: 6

Sequence: 1 gcctcu 6

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 100.0 | 15 | 14 | B0511821 | B0511821 EST619236 |
| 2 | 100.0 | 16 | 14 | B0789979 | B0789979 hage0056A |
| 3 | 100.0 | 19 | 9 | A1569191 | A1569191 tr83112.x |
| 4 | 100.0 | 19 | 17 | A2429998 | A2429998 IM0214F16 |
| 5 | 100.0 | 19 | 17 | A2475079 | A2475079 IM0293B17 |
| 6 | 100.0 | 19 | 17 | A2480905 | A2480905 IM0302N22 |

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 7 | 100.0 | 19 | 17 | A2509071 | A2509071 IM0351A21 |
| 8 | 100.0 | 19 | 17 | A2651803 | A2651803 IM0532N11 |
| 9 | 100.0 | 19 | 17 | A2659603 | A2659603 IM0537N06 |
| 10 | 100.0 | 19 | 17 | A2800646 | A2800646 2M0058H14 |
| 11 | 100.0 | 19 | 17 | A2834038 | A2834038 2M0116H01 |
| 12 | 100.0 | 20 | 14 | B0789776 | B0789776 hage002Aa |
| 13 | 100.0 | 20 | 17 | A2309156 | A2309156 IM0013B09 |
| 14 | 100.0 | 20 | 17 | A2366535 | A2366535 IM0115M15 |
| 15 | 100.0 | 20 | 17 | A2615164 | A2615164 IM0444B24 |
| 16 | 100.0 | 20 | 17 | A2797468 | A2797468 2M0053P09 |
| 17 | 100.0 | 20 | 17 | A2807038 | A2807038 2M0069C06 |
| 18 | 100.0 | 21 | 17 | A2303734 | A2303734 IM0003021 |
| 19 | 100.0 | 21 | 17 | A2317085 | A2317085 IM0003500 |
| 20 | 100.0 | 21 | 17 | A2342282 | A2342282 IM0075H14 |
| 21 | 100.0 | 21 | 17 | A2650869 | A2650869 IM0521G14 |
| 22 | 100.0 | 22 | 9 | A1023917 | A1023917 ow69f05.s |
| 23 | 100.0 | 22 | 9 | A1128425 | A1128425 qc67f09.x |
| 24 | 100.0 | 22 | 9 | A1660937 | A1660937 wf20b06.x |
| 25 | 100.0 | 22 | 17 | A2787102 | A2787102 2M0045F20 |
| 26 | 100.0 | 22 | 17 | A2792724 | A2792724 2M0045F20 |
| 27 | 100.0 | 22 | 17 | A2803482 | A2803482 2M0053I23 |
| 28 | 100.0 | 22 | 17 | TA204A05P | TA204A05P T. Bruce1 |
| 29 | 100.0 | 22 | 17 | TA294D03P | TA294D03P T. Bruce1 |
| 30 | 100.0 | 23 | 17 | A2393604 | A2393604 IM0156C14 |
| 31 | 100.0 | 23 | 17 | A2779667 | A2779667 2M0016C32 |
| 32 | 100.0 | 23 | 17 | A2781980 | A2781980 2M0021N13 |
| 33 | 100.0 | 23 | 17 | A2783504 | A2783504 2M0025L17 |
| 34 | 100.0 | 23 | 17 | A2830526 | A2830526 2M0109M16 |
| 35 | 100.0 | 23 | 17 | A2954682 | A2954682 2M0220P23 |
| 36 | 100.0 | 23 | 17 | A2967993 | A2967993 2M0240J07 |
| 37 | 100.0 | 23 | 17 | BH812514 | BH812514 SALK. 0618 |
| 38 | 100.0 | 24 | 17 | A2443047 | A2443047 IM0237I06 |
| 39 | 100.0 | 24 | 17 | A2463313 | A2463313 IM0272L03 |
| 40 | 100.0 | 24 | 17 | A2505865 | A2505865 IM0346C18 |
| 41 | 100.0 | 24 | 17 | A2780307 | A2780307 2M0017J04 |
| 42 | 100.0 | 24 | 17 | A2787390 | A2787390 2M0033J23 |
| 43 | 100.0 | 24 | 17 | A2789936 | A2789936 2M0038L17 |
| 44 | 100.0 | 24 | 17 | A2812679 | A2812679 2M0079H03 |
| 45 | 100.0 | 24 | 17 | A2936903 | A2936903 2M0193E20 |

ALIGNMENTS

RESULT 1
B0511821
LOCUS
DEFINITION
EST619236 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMH018 5' end, mRNA sequence.

ACCESSION
B0511821
VERSION
B0511821.1
KEYWORDS
SOURCE
ORGANISM
GI:21370690

REFERENCE
AUTHORS
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karamycheva, S.A. set of potato cDNA clones for microarray analyses Generation of a set of potato cDNA clones for microarray analyses Unpublished (2002)
Other ESTs: EST619237
Contact: Robin Buell

TITLE
JOURNAL
COMMENT
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potatoc@tigr.org
This clone is available through the Research Genetics, contact the Research Genetics for further information 1-800-711-6195 or cdna@resgen.com
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1. 15
/organism="Solanum tuberosum"
/cultivar="Kennebec or Bintje"
/db.xref="taxon:4113"
/clone="STM018"
/clone_1lb="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes / tubers, or roots."

BASE COUNT
ORIGIN
3 a 5 c 2 g 5 t

Query Match 100.0%; Score 6; DB 14; Length 15;
Best Local Similarity 83.3%; Pred. No. 5.5e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
Db 9 GCCTCT 14

RESULT 2
LOCUS
DEFINITION
BQ789979 16 bp mRNA linear EST 30-JUL-2002
hage005a12 Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion annosum cDNA clone hage005a12, mRNA sequence.
BQ789979
BQ789979.1 GI:22004941
EST.
Pinus sylvestris/Heterobasidion annosum.
Pinus sylvestris/Heterobasidion annosum
Eukaryota: mixed EST libraries.
1 (bases 1 to 16)
Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
Selected expressed sequence tags of cDNA clones from the interaction of the root rot fungus (Heterobasidion annosum) with seedling roots of Scots pine (Pinus sylvestris)
Unpublished (2001)
Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden
Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer.
Location/Qualifiers
1. 16
/organism="Pinus sylvestris/Heterobasidion annosum"
/db.xref="taxon:169015"
/clone="hage005a12"
/clone_1lb="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
/note="Vector: pT-Adv; Site_1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of Heterobasidion annosum (FP5)."

BASE COUNT
ORIGIN
5 a 3 c 3 g 5 t

Query Match 100.0%; Score 6; DB 14; Length 16;
Best Local Similarity 83.3%; Pred. No. 5.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
Db 7 GCCTCT 12

RESULT 3
LOCUS
DEFINITION
A1569191 19 bp mRNA linear EST 14-MAY-1999
tr83f12.x1 NCI CGAP Pan1 Homo sapiens cDNA clone IMAGE:224943 3' similar to SW:PCCE_HUMAN P28074 PROTEASOME EPSILON CHAIN PRECURSOR
; mRNA sequence.
A1569191
A1569191.1 GI:4532565
EST.
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota: Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 19)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Life Technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: www-bio.11nl.gov/bdrrp/image/image.html

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: www-bio.11nl.gov/bdrrp/image/image.html

Trace considered overall poor quality
Insert length: 1117 Std Error: 0.00
Seq primer: -40UP from Gtlbco
High quality sequence stop: 1
POLYA-No.

FEATURES
source
Location/Qualifiers
1. 19
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:224943"
/clone_1lb="NCI CGAP Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT
ORIGIN
2 a 7 c 5 g 5 t

Query Match 100.0%; Score 6; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6
Db 8 GCCTCT 13

RESULT 4
LOCUS
DEFINITION
A2429998 19 bp DNA linear GSS 03-OCT-2000
IM0214F16F Mouse 10kb plasmid U06CLM library Mus musculus genomic
clone U06CLM0214F16 F, DNA sequence.
A2429998
A2429998.1 GI:10554011
GSS.
ACCESSION
VERSION
KEYWORDS
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weis,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0214 row: F column: 16
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1.19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0214F16"
/clone.lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
5 a 3 c 9 g 2 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 11 GCCTCT 6

RESULT 5
AZ475079/c 19 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0293B17F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
clone UUCG1M0293B17 F, DNA sequence.
ACCESSION AZ475079
VERSION AZ475079.1 GI:10633204
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weis,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0293 row: B column: 17
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1.19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0293B17"
/clone.lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 4 c 9 g 0 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 18 GCCTCT 13

RESULT 6
AZ480905 19 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0302N22R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
clone UUCG1M0302N22 R, DNA sequence.
ACCESSION AZ480905
VERSION AZ480905.1 GI:10642066
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 22
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0302N22"
/clone_lib="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
1 5 c 9 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6,2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||
Db 7 GCCTCT 12

RESULT 7
AZ509071/c 19 bp DNA linear GSS 05-OCT-2000
LOCUS
DEFINITION
1M0351A21R Mouse 10kb plasmid UGC1M library Mus musculus genomic
clone UGC1M0351A21 R, DNA sequence.
ACCESSION
AZ509071 GI:10690387
VERSION
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0351 row: A column: 21
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0351A21"
/clone_lib="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
4 4 c 9 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6,2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
|||||
Db 13 GCCTCT 8

RESULT 8
AZ651803 19 bp DNA linear GSS 14-DEC-2000
LOCUS
DEFINITION
1M0522N11R Mouse 10kb plasmid UGC1M library Mus musculus genomic
clone UGC1M0522N11 R, DNA sequence.
ACCESSION
AZ651803 GI:11787672
VERSION
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.,

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0522 row: N column: 11
Seq primer: CACACGAGAAACAGCATGACCC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0522N11"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[db|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 7 c 5 g 4 t

ORIGIN

Query Match
100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 7 GCCTCT 12

RESULT 9
AZ659603/c 19 bp DNA linear GSS 14-DEC-2000
LOCUS
DEFINITION 1M0537N06F Mouse 10kb plasmid UGGCM library Mus musculus genomic
ACCESSION clone UGGCM0537N06 F, DNA sequence.
AZ659603
VERSION AZ659603.1 GI:11796749
KEYWORDS GSS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.,

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0537 row: N column: 06
Seq primer: CGTTGTAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0537N06"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[db|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
7 a 2 c 7 g 3 t

ORIGIN

Query Match
100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||

Db 12 GCCTCT 7

RESULT 10
AZ800646/c 19 bp DNA linear GSS 16-FEB-2001
LOCUS
DEFINITION 2M0058H14R Mouse 10kb plasmid UGGCM library Mus musculus genomic
ACCESSION clone UGGC2M0058H14 R, DNA sequence.
AZ800646
VERSION AZ800646.1 GI:12952969
KEYWORDS GSS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.,
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Genome Center
Bldg. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0058 row: H column: 14
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0058H14"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g114732114[gb]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 4 c 7 g 5 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 7 GCCTCT 2

RESULT 11
A2834038 19 bp DNA linear GSS 20-FEB-2001
LOCUS 2M0116H01R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCG2M0116H01 R, DNA sequence.
ACCESSION A2834038
VERSION A2834038.1 GI:13003946
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.,
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Genome Center
Bldg. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0116 row: H column: 01
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0116H01"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g114732114[gb]AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 2 c 7 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 6 GCCTCT 1

RESULT 12
BQ789776 20 bp mRNA linear EST 30-JUL-2002
LOCUS hage002a10 Heterobasidion annosum - Scots pine infection stage
DEFINITION (HAGE) subtraction cDNA library plus sylvestrin/Heterobasidion
annosum cDNA clone hage002a10, mRNA sequence.
ACCESSION BQ789776
VERSION BQ789776.1 GI:22004738
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 20)

AUTHORS Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
TITLE Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (*Heterobasidion annosum*) with seedling roots of Scots pine (*Pinus sylvestris*)
JOURNAL Unpublished (2001)
COMMENT Contact: Fred O. Asiegbu
 Dept. of Forest Mycology & Pathology
 Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden
 Tel: +46 18 67 15 98
 Fax: +46 18 30 92 45
 Email: Fred.Asiegbu@mykopat.slu.se
 Seq primer: T7 primer.

FEATURES
 source location/Qualifiers
 1..20
 /organism="Pinus sylvestris/Heterobasidion annosum"
 /db_xref="taxon:169015"
 /clone="hage002a10"
 /clone.lib="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
 /dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
 /note="Vector: PT-Adv; Site: 1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of *Heterobasidion annosum* (Fp5)."

BASE COUNT
 4 a 5 c 3 g 7 t 1 others

ORIGIN

Query Match 100.0%; Score 6; DB 14; Length 20;
 Best Local Similarity 83.3%; Pred. No. 6.4e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
 |||||:
 Db 12 GCCTCT 17

RESULT 13
AZ309156 20 bp DNA linear GSS 29-SEP-2000
LOCUS IM0013B09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION Clone UUGC1M0013B09 F, DNA sequence.
ACCESSION AZ309156
VERSION AZ309156.1 GI:10349862
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weis, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0013 row: B column: 09
 Seq primer: CCGTCAAAACGACGCCACT
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers
 1..20
 source

BASE COUNT
 2 a 11 c 4 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 20;
 Best Local Similarity 83.3%; Pred. No. 6.4e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCT 6
 |||||:
 Db 3 GCCTCT 8

RESULT 14
AZ366535 20 bp DNA linear GSS 02-OCT-2000
LOCUS IM0115M15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION Clone UUGC1M0115M15 R, DNA sequence.
ACCESSION AZ366535
VERSION AZ366535.1 GI:10480235
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weis, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0115 row: M column: 15
 Seq primer: CACACAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers
 1..20
 source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0115M15"
/clone_1lb="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      5 a      6 c      3 g      6 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GCCTCU 6
       11111:
Db      14 GCCTCT 19

```

```

RESULT 15
LOCUS      A2615164      20 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION 1M0444B24F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
ACCESSION  A2615164
VERSION    A2615164.1 GI:11737354
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE  1 (bases 1 to 20)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
            and Wright,D., Weis,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
JOURNAL    University of Utah Genome Center
COMMENT    Contact: Robert B. Weiss
            University of Utah
            Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177

```

```

FEATURES
source     Email: cdunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0444 row: B column: 24
            Seq primer: CGTGTGAACGACGCGCCAGT
            Class: plasmid ends
            High quality sequence stop: 20.
            Location/Qualifiers
            1..20

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0444B24"
/clone_1lb="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      3 a      7 c      5 g      5 t
ORIGIN

```

```

Query Match      100.0%; Score 6; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GCCTCU 6
       11111:
Db      7 GCCTCT 12

```

```

Search completed: May 21, 2003, 06:26:03
Job time : 1094 secs

```

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 04:39:53 ; Search time 238 Seconds

(without alignments)
33.289 Million cell updates/sec

Title: US-09-936-146-9

Sequence: 1 gcctcu 6

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_MA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1 | 6 | 100.0 | 9 | US-09-990-186-588 | Sequence 588, App |
| 2 | 6 | 100.0 | 9 | US-09-990-186-2084 | Sequence 2084, Ap |
| 3 | 6 | 100.0 | 9 | US-09-990-186-2086 | Sequence 2086, Ap |
| 4 | 6 | 100.0 | 9 | US-09-990-186-2219 | Sequence 2219, Ap |
| 5 | 6 | 100.0 | 9 | US-09-990-186-2280 | Sequence 2280, Ap |
| 6 | 6 | 100.0 | 9 | US-09-990-186-2471 | Sequence 2471, Ap |
| 7 | 6 | 100.0 | 9 | US-09-990-186-2473 | Sequence 2473, Ap |
| 8 | 6 | 100.0 | 9 | US-09-989-789-588 | Sequence 588, App |
| 9 | 6 | 100.0 | 9 | US-09-989-789-2084 | Sequence 2084, Ap |
| 10 | 6 | 100.0 | 9 | US-09-989-789-2086 | Sequence 2086, Ap |
| 11 | 6 | 100.0 | 9 | US-09-989-789-2219 | Sequence 2219, Ap |
| 12 | 6 | 100.0 | 9 | US-09-989-789-2280 | Sequence 2280, Ap |
| 13 | 6 | 100.0 | 9 | US-09-989-789-2471 | Sequence 2471, Ap |
| 14 | 6 | 100.0 | 9 | US-09-989-789-2473 | Sequence 2473, Ap |
| 15 | 6 | 100.0 | 10 | US-09-990-186-617 | Sequence 617, App |
| 16 | 6 | 100.0 | 10 | US-09-990-186-1279 | Sequence 1279, Ap |
| 17 | 6 | 100.0 | 10 | US-09-990-186-1308 | Sequence 1308, Ap |
| 18 | 6 | 100.0 | 10 | US-09-990-186-1313 | Sequence 1313, Ap |
| 19 | 6 | 100.0 | 10 | US-09-989-789-617 | Sequence 617, App |

| | | | | | |
|----|---|-------|----|---------------------|--------------------|
| 20 | 6 | 100.0 | 10 | US-09-989-789-1279 | Sequence 1279, Ap |
| 21 | 6 | 100.0 | 10 | US-09-989-789-1308 | Sequence 1308, Ap |
| 22 | 6 | 100.0 | 10 | US-09-989-789-1313 | Sequence 1313, Ap |
| 23 | 6 | 100.0 | 10 | US-10-033-145-153 | Sequence 153, App |
| 24 | 6 | 100.0 | 10 | US-10-033-145-174 | Sequence 174, App |
| 25 | 6 | 100.0 | 10 | US-10-033-145-355 | Sequence 355, App |
| 26 | 6 | 100.0 | 10 | US-10-033-145-566 | Sequence 566, App |
| 27 | 6 | 100.0 | 10 | US-10-033-145-968 | Sequence 968, App |
| 28 | 6 | 100.0 | 10 | US-10-033-145-1007 | Sequence 1007, App |
| 29 | 6 | 100.0 | 10 | US-10-033-145-1120 | Sequence 1120, Ap |
| 30 | 6 | 100.0 | 10 | US-10-033-145-1286 | Sequence 1286, Ap |
| 31 | 6 | 100.0 | 10 | US-10-033-145-1609 | Sequence 1609, Ap |
| 32 | 6 | 100.0 | 11 | US-09-249-155-83 | Sequence 83, Appl |
| 33 | 6 | 100.0 | 11 | US-09-249-155-207 | Sequence 207, App |
| 34 | 6 | 100.0 | 11 | US-09-249-155-239 | Sequence 239, App |
| 35 | 6 | 100.0 | 11 | US-09-249-155-239 | Sequence 239, Appl |
| 36 | 6 | 100.0 | 12 | US-09-761-116-1 | Sequence 1, Appl |
| 37 | 6 | 100.0 | 12 | US-09-384-472-18 | Sequence 18, Appl |
| 38 | 6 | 100.0 | 12 | US-09-384-472-21 | Sequence 21, Appl |
| 39 | 6 | 100.0 | 12 | US-08-591-486B-102 | Sequence 102, App |
| 40 | 6 | 100.0 | 14 | US-09-978-600-198 | Sequence 198, App |
| 41 | 6 | 100.0 | 14 | US-09-998-027-26 | Sequence 26, Appl |
| 42 | 6 | 100.0 | 14 | US-09-504-231A-1431 | Sequence 1431, Ap |
| 43 | 6 | 100.0 | 14 | US-09-274-553D-1431 | Sequence 1431, Ap |
| 44 | 6 | 100.0 | 15 | US-10-056-414-119 | Sequence 119, App |
| 45 | 6 | 100.0 | 15 | | |

ALIGNMENTS

RESULT 1
US-09-990-186-588/c
Sequence 588, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 588
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-588

Query Match
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 8 GCCTCT 3

RESULT 2
US-09-990-186-2084/c
Sequence 2084, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT FILING DATE: 2001-11-20

NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2084
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2084

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 3
US-09-990-186-2086/c
Sequence 2086, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2086
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2086

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 4
US-09-990-186-2219/c
Sequence 2219, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2219
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2219

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 5
US-09-990-186-2280/c
Sequence 2280, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2280
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2280

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 6
US-09-990-186-2471/c
Sequence 2471, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: LIU, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990,186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2471
LENGTH: 9
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2471

Query Match 100.0%; Score 6; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
DB 8 GCCTCT 3

RESULT 7
US-09-990-186-2473/c

```

; Sequence 2473, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-990-186-2473

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GCCTCU 6
    |||||
Db 8 GCCTCT 3

```

```

RESULT 8
US-09-989-789-588/c
; Sequence 588, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 588
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-588

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GCCTCU 6
    |||||
Db 8 GCCTCT 3

```

```

RESULT 9
US-09-989-789-2084/c
; Sequence 2084, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0

```

```

; SEQ ID NO 2084
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2084

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GCCTCU 6
    |||||
Db 8 GCCTCT 3

```

```

RESULT 10
US-09-989-789-2086/c
; Sequence 2086, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2086
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2086

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GCCTCU 6
    |||||
Db 8 GCCTCT 3

```

```

RESULT 11
US-09-989-789-2219/c
; Sequence 2219, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2219
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
US-09-989-789-2219

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 12
US-09-989-789-2280/c
; Sequence 2280, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2280
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2280

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 13
US-09-989-789-2471/c
; Sequence 2471, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2471
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2471

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 14
US-09-989-789-2473/c
; Sequence 2473, Application US/09989789
; Patent No. US20020063379A1

; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2473
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-989-789-2473

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.5e+08;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

RESULT 15
US-09-990-186-617/c
; Sequence 617, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example target
; OTHER INFORMATION: DNA
US-09-990-186-617

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 8 GCCTCT 3

Search completed: May 21, 2003, 06:44:44
Job time : 238 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: May 21, 2003, 05:13:24 ; Search time 331.5 Seconds

(without alignments)
526.748 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6

Sequence: 1 atgacg 6

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 6 | 100.0 | 8 | 6 | AX359003 | AX359003 Sequence |
| 2 | 6 | 100.0 | 10 | 6 | AR104790 | AR104790 Sequence |
| 3 | 6 | 100.0 | 10 | 6 | AX113032 | AX113032 Sequence |
| 4 | 6 | 100.0 | 10 | 6 | AX152230 | AX152230 Sequence |
| 5 | 6 | 100.0 | 10 | 6 | AX152244 | AX152244 Sequence |
| 6 | 6 | 100.0 | 10 | 6 | AX152248 | AX152248 Sequence |
| 7 | 6 | 100.0 | 10 | 6 | AX152409 | AX152409 Sequence |
| 8 | 6 | 100.0 | 10 | 6 | AX152854 | AX152854 Sequence |
| 9 | 6 | 100.0 | 10 | 6 | AX152855 | AX152855 Sequence |
| 10 | 6 | 100.0 | 10 | 6 | AX153355 | AX153355 Sequence |
| 11 | 6 | 100.0 | 10 | 6 | AX239915 | AX239915 Sequence |
| 12 | 6 | 100.0 | 10 | 6 | AX301406 | AX301406 Sequence |
| 13 | 6 | 100.0 | 10 | 6 | AX301520 | AX301520 Sequence |
| 14 | 6 | 100.0 | 10 | 6 | AX301566 | AX301566 Sequence |
| 15 | 6 | 100.0 | 10 | 6 | AX301578 | AX301578 Sequence |
| 16 | 6 | 100.0 | 11 | 6 | AX377149 | AX377149 Sequence |
| 17 | 6 | 100.0 | 11 | 6 | AX022911 | AX022911 Sequence |
| 18 | 6 | 100.0 | 11 | 6 | AX022930 | AX022930 Sequence |
| 19 | 6 | 100.0 | 11 | 6 | AX022949 | AX022949 Sequence |
| 20 | 6 | 100.0 | 11 | 6 | AX030499 | AX030499 Sequence |
| 21 | 6 | 100.0 | 11 | 6 | AX030518 | AX030518 Sequence |
| 22 | 6 | 100.0 | 11 | 6 | AX030537 | AX030537 Sequence |
| 23 | 6 | 100.0 | 11 | 6 | AX471026 | AX471026 Sequence |
| 24 | 6 | 100.0 | 11 | 6 | AX471139 | AX471139 Sequence |
| 25 | 6 | 100.0 | 11 | 6 | AX471204 | AX471204 Sequence |
| 26 | 6 | 100.0 | 11 | 6 | AX471618 | AX471618 Sequence |
| 27 | 6 | 100.0 | 11 | 6 | AX471768 | AX471768 Sequence |
| 28 | 6 | 100.0 | 11 | 6 | AX472198 | AX472198 Sequence |
| 29 | 6 | 100.0 | 11 | 6 | I28557 | I28557 Sequence 10 |
| 30 | 6 | 100.0 | 11 | 6 | I56092 | I56092 Sequence 3 |
| 31 | 6 | 100.0 | 11 | 6 | I58719 | I58719 Sequence 10 |
| 32 | 6 | 100.0 | 12 | 6 | AR058577 | AR058577 Sequence |
| 33 | 6 | 100.0 | 12 | 6 | AR058578 | AR058578 Sequence |
| 34 | 6 | 100.0 | 12 | 6 | AR167775 | AR167775 Sequence |
| 35 | 6 | 100.0 | 12 | 6 | AR193085 | AR193085 Sequence |
| 36 | 6 | 100.0 | 12 | 6 | AX464720 | AX464720 Sequence |
| 37 | 6 | 100.0 | 12 | 6 | E29659 | E29659 Method for |
| 38 | 6 | 100.0 | 12 | 6 | E38765 | E38765 Method for |
| 39 | 6 | 100.0 | 12 | 6 | E64191 | E64191 Method for |
| 40 | 6 | 100.0 | 13 | 6 | I04302 | I04302 Sequence 15 |
| 41 | 6 | 100.0 | 13 | 6 | AR069833 | AR069833 Sequence |
| 42 | 6 | 100.0 | 13 | 6 | AR149120 | AR149120 Sequence |
| 43 | 6 | 100.0 | 13 | 6 | AX048318 | AX048318 Sequence |
| 44 | 6 | 100.0 | 13 | 6 | AX266958 | AX266958 Sequence |
| 45 | 6 | 100.0 | 13 | 6 | AX300869 | AX300869 Sequence |

ALIGNMENTS

RESULT 1
AX359003
LOCUS AX359003
DEFINITION Sequence 10 from Patent WO0183737.
ACCESSION AX359003
VERSION AX359003.1 GI:18675402
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
Nagy, E., Tuboly, T. and Nagy, M.
Porcine adenovirus vaccine
Patent: WO 0183737-A 10 08-NOV-2001;
UNIVERSITY OF GUELPH (CA)

FEATURES Location/Qualifiers
source 1..8
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Fragment"

BASE COUNT 3 a 1 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 8;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 1 ATGACT 6

RESULT 2
ARI04790 10 bp DNA linear PAT 14-FEB-2001
LOCUS
DEFINITION Sequence 87 from patent US 6093811.
ACCESSION ARI04790
VERSION ARI04790.1 GI:12817498
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Bennett,C.Frank. and Mirabelli,C.K.
TITLE Oligonucleotide modulation of cell adhesion
JOURNAL Patent: US 6093811-A 87 25-JUL-2000;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 4 a 2 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 5 ATGACT 10

RESULT 3
AX113032 10 bp DNA linear PAT 01-MAY-2001
LOCUS
DEFINITION Sequence 79 from Patent WO0127267.
ACCESSION AX113032
VERSION AX113032.1 GI:13939467
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 10)
AUTHORS Adams,E., Waldmann,H., Cobbold,S. and Zelenika,D.
TITLE Genes differentially expressed in trl cells and their use in the
manufacture of immunoregulatory compositions
JOURNAL Patent: WO 0127267-A 79 19-APR-2001;
FEATURES Location/Qualifiers
source 1..10
/organism="Mus sp."
/db_xref="taxon:10095"

BASE COUNT 4 a 1 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 1 ATGACT 6

RESULT 4
AX152230/C 10 bp DNA linear PAT 22-JUN-2001
LOCUS
DEFINITION Sequence 145 from Patent WO0138577.
ACCESSION AX152230
VERSION AX152230.1 GI:14533881
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 145 31-MAY-2001;
FEATURES Location/Qualifiers
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 3 a 2 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 9 ATGACT 4

RESULT 5
AX152244/C 10 bp DNA linear PAT 22-JUN-2001
LOCUS
DEFINITION Sequence 159 from Patent WO0138577.
ACCESSION AX152244
VERSION AX152244.1 GI:14533895
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 10)
AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcripts
JOURNAL Patent: WO 0138577-A 159 31-MAY-2001;
FEATURES Location/Qualifiers
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 3 a 2 c 2 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 10 ATGACT 5

RESULT 6
AX152408 10 bp DNA linear PAT 22-JUN-2001
LOCUS

DEFINITION Sequence 323 from Patent WO0138577.
ACCESSION AX152408
VERSION AX152408.1 GI:14534059
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 10)
TITLE Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
JOURNAL Human transcriptomes
Patent: WO 0138577-A 323 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 1 ATGACT 6

RESULT 7
AX152409 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 324 from Patent WO0138577.
DEFINITION AX152409
ACCESSION AX152409.1 GI:14534060
VERSION
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 10)
TITLE Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
JOURNAL Human transcriptomes
Patent: WO 0138577-A 324 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 1 ATGACT 6

RESULT 8
AX152854/c 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 769 from Patent WO0138577.
DEFINITION AX152854
ACCESSION AX152854
VERSION AX152854.1 GI:14534505
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 10)

AUTHORS Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
TITLE Human transcriptomes
JOURNAL Patent: WO 0138577-A 769 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 7 ATGACT 2

RESULT 9
AX152855 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 770 from Patent WO0138577.
DEFINITION AX152855
ACCESSION AX152855
VERSION AX152855.1 GI:14534506
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 10)
TITLE Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
JOURNAL Human transcriptomes
Patent: WO 0138577-A 770 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 2 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 7 ATGACT 2

RESULT 10
AX153355 10 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 1270 from Patent WO0138577.
DEFINITION AX153355
ACCESSION AX153355
VERSION AX153355.1 GI:14535006
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 10)
TITLE Velculescu,V.E., Vogelstein,B. and Kinzler,K.W.
JOURNAL Human transcriptomes
Patent: WO 0138577-A 1270 31-MAY-2001;
The Johns Hopkins University (US)
FEATURES
source 1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2 a 1 c 4 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
8 ATGACT 3

RESULT 11
AX239915 10 bp DNA linear PAT 26-SEP-2001
LOCUS Sequence 42 from Patent WO0164958.
DEFINITION AX239915
ACCESSION AX239915
VERSION AX239915.1 GI:15797517
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 10)
AUTHORS Dempcy,R.O., Gall,A.A., Likhov,S.G., Afonina,I.A., Singer,M.J.,
Kutyavina,I.V. and Vermeulen,N.M.
TITLE Modified oligonucleotides for mismatch discrimination
JOURNAL Patent: WO 0164958-A 42 07-SEP-2001;
Epoch Biosciences, Inc. (US)
FEATURES
source location/Qualifiers
1..10
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="duplex complement 7"

BASE COUNT 4 a 2 c 2 g 2 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
5 ATGACT 10

RESULT 12
AX301406 10 bp DNA linear PAT 30-NOV-2001
LOCUS AX301406
DEFINITION Sequence 120 from Patent WO0185941.
ACCESSION AX301406
VERSION AX301406.1 GI:17382489
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
TITLE Versteeg,R. and Caron,H.N.
JOURNAL MYC targets
Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)
FEATURES
source location/Qualifiers
1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 4 a 1 c 2 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
1 ATGACU 6

DB 8 ATGACT 3

RESULT 13
AX301520 10 bp DNA linear PAT 30-NOV-2001
LOCUS AX301520
DEFINITION Sequence 234 from Patent WO0185941.
ACCESSION AX301520
VERSION AX301520.1 GI:17382603
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
TITLE Versteeg,R. and Caron,H.N.
JOURNAL MYC targets
Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)
FEATURES
source location/Qualifiers
1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 4 a 1 c 1 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
2 ATGACT 7

RESULT 14
AX301566 10 bp DNA linear PAT 30-NOV-2001
LOCUS AX301566
DEFINITION Sequence 280 from Patent WO0185941.
ACCESSION AX301566
VERSION AX301566.1 GI:17382649
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
TITLE Versteeg,R. and Caron,H.N.
JOURNAL MYC targets
Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)
FEATURES
source location/Qualifiers
1..10
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 4 a 1 c 1 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
2 ATGACT 7

RESULT 15
AX301578 10 bp DNA linear PAT 30-NOV-2001
LOCUS AX301578
DEFINITION Sequence 292 from Patent WO0185941.
ACCESSION AX301578
VERSION AX301578.1 GI:17382661

KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Versteeg,R. and Caron,H.N.
TITLE Myc targets
JOURNAL Patent: WO 0185941-A 292 15-NOV-2001;
Academisch Ziekenhuis bij de Universiteit van Amsterdam (NL)
FEATURES
source 1.
10
/Organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 1 c 2 g 3 t
ORIGIN
Query Match 100.0%; Score 6; DB 6; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.4e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACU 6
|||||:
Db 8 ATGACT 3

Search completed: May 21, 2003, 07:17:17
Job time : 334.5 secs


```

RESULT 1
AAAB0753/c
ID AAA80753 standard; DNA; 8 BP.
XX
AC
XX
AAAB0753;
XX
24-NOV-2000 (first entry)
XX
DE A. thaliana primer walking octamer SEQ ID NO: 66.
XX
XX Primer walking; octamer; primer; DNA sequencing; PCR; ss.
XX
OS Arabidopsis thaliana.
XX
XX US6083695-A.
XX
XX PN
XX
XX 04-JUL-2000.
XX
XX
XX
XX 21-MAY-1997; 97US-0859954.
XX
XX
XX 15-APR-1996; 96US-0632782.
XX
XX (UYHO-) UNIV HOUSTON.
XX
XX PA (HARD/) HARDIN S H.
XX
XX
XX PI Hardin PE, Hardin SH, Homayouni R;
XX
XX WPI; 2000-474852/41.
XX
XX
XX Sequencing an unknown DNA molecule for the polymerase chain reaction
XX PT and other primer processes comprises primer walking of octamer
XX PT oligonucleotides -

```

XX Example 8; Column 57-58; 161pp; English.
PS

CC This invention describes a novel method for sequencing an unknown DNA
CC molecule which comprises selecting a library primer from an octamer
CC oligonucleotide library consisting of 48 8-bp sequences and
CC corresponding complementary sequences, where the library primer is
CC complementary to a known sequence adjacent to the unknown sequence or
CC is complementary to a sequence in a known extension product. The method
CC is useful for DNA nucleotide sequencing, in PCR, and in other processes
CC which make use of primers. The octamers are used to identify coding
CC sequences. Primer walking using the octamer libraries is advantageous
CC over other sequencing methods because it does not require multiple
CC cloning steps nor subsequent template preparations, and it is a
CC directed and methodical approach. AA080688-A01253 represent the octamer
CC primers used in the primer walking method of the invention.

SQ Sequence 8 BP; 2 A; 1 C; 3 G; 2 T; 0 other;

| | | | | |
|-----------------------|--------|--------------|----------|---------------|
| Query Match | 100.0% | Score 6: | DB 21; | Length 8; |
| Best Local Similarity | 83.3% | Pred. No. | 2.7e+08; | |
| Matches | 5; | Conservative | 1; | Mismatches 0; |
| | | | | Gaps 0; |

| | | | |
|----|---|--------|---|
| QY | 1 | ATGACU | 6 |
| | | : | |
| Db | 8 | ATGACT | 3 |

RESULT 2
AAA80845

AC AAA80845 ;

DT 24-NOV-2000 (first entry)

DE A. thaliana primer walking octamer SEQ ID NO: 158.

KW Primer walking; octamer; primer; DNA sequencing; PCR; ss

OS *Arabidopsis thaliana*.

PN US6083695-A.

PD 04-JUL-2000

PF 21-MAY-1997; 97US-0859954.

PR 15-APR-1996; 96US-0632782.

PA (UYHO-) UNIV HOUSTON.

XX
XX

[illegible]

XX

PT Sequencing an unknown DNA molecule for the polymerase chain reaction
PT and other primer processes comprises primer walking of octamer
PT oligonucleotides -

PS Example 8; Column 105-106; 161pp; English.

This invention describes a novel method for sequencing an unknown DNA molecule which comprises selecting a library primer from an octamer oligonucleotide library consisting of 48 8-bp sequences and corresponding complementary sequences, where the library primer is complementary to a known sequence adjacent to the unknown sequence or is complementary to a sequence in a known extension product. The method is useful for DNA nucleotide sequencing, in PCR, and in other processes which make use of primers. The octamers are used to identify coding sequences. Primer walking using the octamer libraries is advantageous over other sequencing methods because it does not require multiple

CC cloning steps nor subsequent template preparations, and it is a
CC directed and methodical approach. AAA80688-A81253 represent the octamer
CC primers used in the primer walking method of the invention.

Sequence 8 BP; 2 A; 2 C; 2 G; 2 T; 0 other;

| | | | | |
|-----------------------|--------|--------------|----------|---------------|
| Query Match | 100.0% | Score 6: | DB 21: | length 8: |
| Best Local Similarity | 83.3% | Pred. No. | 2.7e+08: | |
| Matches | 5: | Conservative | 1: | Mismatches 0: |
| | | | | Indels 0: |
| | | | | Gaps 0: |

| | | | |
|----|---|--------|---|
| Qy | 1 | ATGACU | 6 |
| | | : | |
| Db | 2 | ATGACT | 7 |

RESULT 3
ABK87314
ID ABK87314 standard; DNA; 9 BP.

AC ABK87314;

DT 24-SEP-2002 (first entry)

GCN4 yeast recognition sequence #1.

KW Nucleic acid detection; GCN4; lambda repressor; ss; yeast

05 *Saccharomyces*.

PN WO200244326-A2

06-JUN-2002.

PF 26-NOV-2001; 2001WO-US44215.

PR 30-NOV-2000; 2000US-0728574.

PA (STRA-) STRATAGENE
xy

PI Sørge JA, Whalen AM;
xy

DR WPI; 2002-508503/54.

PT Detecting/measuring target nucleic acid, by forming cleavage structure
PT by incubating target nucleic acid with probe having binding moiety,
PT cleaving structure to release nucleic acid and detecting released
PT fragments -

PS Disclosure; Page 75; 157pp; English.

CC This invention relates to a novel method for detecting/measuring a
CC target nucleic acid. The method comprises forming a cleavage structure
CC by incubating the target sequence with a probe comprising a binding
CC moiety and a secondary structure that changes upon binding of the probe
CC to the target, cleaving the cleavage structure to release a nucleic
CC acid fragment, and detecting and/or measuring the fragment captured by
CC binding of the binding moiety to a capture element on a solid support.
CC The method of the invention is useful for detecting or measuring a
CC target nucleic acid and are useful for generating a signal indicative of
CC the presence of the target nucleic acid in a sample. Another method of
CC the invention is useful for simultaneously forming a cleavage structure,
CC amplifying the target nucleic acid in a sample and cleaving the cleavage
CC structure. The method does not require multiple steps, subsequent
CC amplification process, and allows for concurrent amplification and
CC detection of target nucleic acid in a sample. The present sequence
CC represents the yeast GCN4 recognition sequence shown in the
CC specification.

sq` Sequence 9 BP; 3 A; 2 C; 1 G; 3 T; 0 other;

| | | | | |
|-------------|-----------|--------------|-------------------|---------------|
| Query Match | 100.0% | Score 6: | DB 24; | Length 9; |
| Best Local | Similarly | 83.3%; | Pred.No. 2.4e+08; | |
| Matches | 5; | Conservative | 1; | Mismatches 0; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

OY 1 ATGACU 6
 OS |||||:
 PN 1 ATGACT 6
 DB

RESULT 4
 AAQ96636
 ID AAQ96636 standard; DNA; 10 BP.

XX AC AAQ96636;

XX DT 22-MAR-1996 (first entry)

XX DE HIV-1 NL4-3 nef gene nucleotide deletion 231.

XX KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX OS Human immunodeficiency virus type 1.

XX PN WO9521912-A1.

XX PD 17-AUG-1995.

XX PF 14-FEB-1995; 95WO-AU00063.

XX PR 23-DEC-1994; 94AU-0000284.

XX PR 14-FEB-1994; 94AU-0003864.

XX PR 21-FEB-1994; 94AU-0004002.

XX PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX DR WPI; 1995-293115/38.

XX PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

XX PT or LTR region - can be used in a vaccine to inhibit/reduce

XX PT productive infection in an individual by a pathogenic strain

XX PS Claim 13; Page 191; 301pp; English.

XX CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

XX CC or more deancuclotides (AAQ96406-097018) from the nef gene and/or

XX CC 1 or more deancuclotides (AAQ97019-097166) from the LTR region; the

XX CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene

XX CC (AAQ96141). The resulting avirulent HIV strains are still capable of

XX CC inducing an immune response in humans, and enable the generation of

XX CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX SQ Sequence 10 BP; 4 A; 3 C; 1 G; 2 T; 0 other;

XX OY Query Match 100.0%; Score 6; DB 16; Length 10;

XX ID Best Local Similarity 83.3%; Pred. No. 1.7e+05;

XX AC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX DT 1 ATGACU 6

XX DB 5 ATGACT 10

RESULT 5
 AAQ96637
 ID AAQ96637 standard; DNA; 10 BP.

XX AC AAQ96637;

XX DT 22-MAR-1996 (first entry)

XX DE HIV-1 NL4-3 nef gene nucleotide deletion 232.

XX KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX OS Human immunodeficiency virus type 1.
 XX PN WO9521912-A1.
 XX PD 17-AUG-1995.
 XX PF 14-FEB-1995; 95WO-AU00063.
 XX PR 23-DEC-1994; 94AU-0000284.
 XX PR 14-FEB-1994; 94AU-0003864.
 XX PR 21-FEB-1994; 94AU-0004002.

XX PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
 XX PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX DR WPI; 1995-293115/38.

XX PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene

XX PT or LTR region - can be used in a vaccine to inhibit/reduce

XX PT productive infection in an individual by a pathogenic strain

XX PS Claim 13; Page 191; 301pp; English.

XX CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1

XX CC or more deancuclotides (AAQ96406-097018) from the nef gene and/or

XX CC 1 or more deancuclotides (AAQ97019-097166) from the LTR region; the

XX CC sequence of AAQ96406 corresponds to nucleotides 1-10 of the nef gene

XX CC (AAQ96141). The resulting avirulent HIV strains are still capable of

XX CC inducing an immune response in humans, and enable the generation of

XX CC therapeutic, diagnostic and targeting agents against HIV-1 infection.

XX SQ Sequence 10 BP; 3 A; 3 C; 1 G; 3 T; 0 other;

XX OY Query Match 100.0%; Score 6; DB 16; Length 10;

XX ID Best Local Similarity 83.3%; Pred. No. 1.7e+05;

XX AC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX DT 1 ATGACU 6

XX DB 4 ATGACT 9

RESULT 6
 AAQ96638
 ID AAQ96638 standard; DNA; 10 BP.

XX AC AAQ96638;

XX DT 22-MAR-1996 (first entry)

XX DE HIV-1 NL4-3 nef gene nucleotide deletion 233.

XX KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX OS Human immunodeficiency virus type 1.

XX PN WO9521912-A1.

XX PD 17-AUG-1995.

XX PF 14-FEB-1995; 95WO-AU00063.

XX PR 23-DEC-1994; 94AU-0000284.

XX PR 14-FEB-1994; 94AU-0003864.

XX PR 21-FEB-1994; 94AU-0004002.

XX PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.

XX PI (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;

XX WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 191; 301pp; English.
XX
CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decaunucleotides (AA096406-Q97018) from the nef gene and/or
CC 1 or more decaunucleotides (AA097019-Q97166) from the LTR region; the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
SQ Sequence 10 BP; 4 A; 2 C; 1 G; 3 T; 0 other;
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. NO. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGACU 6
DB 3 ATGACT 8
RESULT 7
AA096639 standard; DNA: 10 BP.
ID AA096639 standard; DNA: 10 BP.
XX
AC AA096639;
XX
DT 22-MAR-1996 (first entry)
XX
DE HIV-1 NL4-3 nef gene nucleotide deletion 234.
XX
KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO9521912-A1.
XX
PD 17-AUG-1995.
XX
PE 14-FEB-1995; 95WO-AU00063.
XX
PR 23-DEC-1994; 94AU-0000284.
PR 14-FEB-1994; 94AU-0003864.
PR 21-FEB-1994; 94AU-0004002.
XX
PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI: 1995-293115/38.
XX
DR WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 191; 301pp; English.
XX
CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decaunucleotides (AA096406-Q97018) from the nef gene and/or
CC 1 or more decaunucleotides (AA097019-Q97166) from the LTR region; the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX

SQ Sequence 10 BP; 4 A; 2 C; 1 G; 3 T; 0 other;
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. NO. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGACU 6
DB 2 ATGACT 7
RESULT 8
AA096640 standard; DNA: 10 BP.
ID AA096640 standard; DNA: 10 BP.
XX
AC AA096640;
XX
DT 22-MAR-1996 (first entry)
XX
DE HIV-1 NL4-3 nef gene nucleotide deletion 235.
XX
KM HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO9521912-A1.
XX
PD 17-AUG-1995.
XX
PE 14-FEB-1995; 95WO-AU00063.
XX
PR 23-DEC-1994; 94AU-0000284.
PR 14-FEB-1994; 94AU-0003864.
PR 21-FEB-1994; 94AU-0004002.
XX
PA (AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
PI Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;
PI WPI: 1995-293115/38.
XX
DR WPI: 1995-293115/38.
XX
XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene
PT or LTR region - can be used in a vaccine to inhibit/reduce
PT productive infection in an individual by a pathogenic strain
XX
XX Claim 13; Page 191; 301pp; English.
XX
CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1
CC or more decaunucleotides (AA096406-Q97018) from the nef gene and/or
CC 1 or more decaunucleotides (AA097019-Q97166) from the LTR region; the
CC sequence of AA096406 corresponds to nucleotides 1-10 of the nef gene
CC (AA096141). The resulting avirulent HIV strains are still capable of
CC inducing an immune response in humans, and enable the generation of
CC therapeutic, diagnostic and targeting agents against HIV-1 infection.
XX
SQ Sequence 10 BP; 4 A; 2 C; 1 G; 3 T; 0 other;
Query Match 100.0%; Score 6; DB 16; Length 10;
Best Local Similarity 83.3%; Pred. NO. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGACU 6
DB 1 ATGACT 6
RESULT 9
AAV50231/C
ID AAV50231 standard; DNA: 10 BP.
XX
AC AAV50231;
XX

DT 21-OCT-1998 (first entry)
 XX Yeast tag for additional NORF chromosome 15 tag position 32081.
 DE
 XX
 XX Yeast; Saccharomyces cerevisiae; transcriptome; cell cycle;
 KM regulation; eukaryotic cell; antifungal; SAGE tag; gene expression;
 KW serial analysis of gene expression; probe; ss.
 XX
 OS Saccharomyces cerevisiae.
 OS Synthetic.
 XX
 PN WO9832847-A2.
 XX
 PD 30-JUL-1998.
 XX
 PF 22-JAN-1998: 98WO-US01216.
 XX
 PR 23-JAN-1997: 97US-0035917.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 XX
 PI Kinzler KW, Velculescu VE, Vogelstein B;
 XX
 DR WPI: 1998-427943/36.
 XX
 PT Yeast transcriptome - useful for modulating eukaryotic cell, for
 PT screening antifungal agents, and for identifying genes in cell cycle
 PT progression
 XX
 PS Claim 1; Page 25; 44pp; English.
 XX
 CC Yeast transcriptome is encoded by a DNA molecule comprising a yeast
 CC gene involved in cell cycle progression selected from the group of
 CC nonannotated ORF (NORF) genes. SAGE (serial analysis gene expression)
 CC tags for highly expressed genes and NORF genes are given in AAV50051 to
 CC AAV50345. The present invention describes: (1) a method of using yeast
 CC genes to modulate the cell cycle which comprises administering to a cell
 CC an isolated DNA molecule comprising a yeast gene which is involved in
 CC cell cycle progression selected from differentially expressed genes
 CC (SAGE tags given in AAV50051 to AAV50345); (2) a method for screening
 CC candidate antifungal drugs which comprises contacting a test substance
 CC with a yeast cell and monitoring expression of a yeast gene which is
 CC involved in cell cycle progression; (3) a method of identifying human
 CC genes which are involved in cell cycle progression which comprises
 CC hybridizing a probe comprising at least 10 contiguous nucleotides of a
 CC yeast gene which is differentially expressed between at least 2 phases
 CC selected from the log phase, the S phase and the G2/M phase; and (4) a
 CC probe for ascertaining the phase in the cell cycle, where the probe
 CC comprises at least 14 contiguous nucleotides of a NORF gene (SAGE tags
 CC given in AAV50051 to AAV50345), or as an array of probes on a solid
 CC support.
 CC
 SO Sequence 10 BP; 5 A; 1 C; 1 G; 3 T; 0 other:
 Query Match 100.0%; Score 6; DB 19; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 ID AAX18647
 XX AAX18647 standard; DNA; 10 BP.
 AC AAX18647;
 XX
 DT 06-MAY-1999 (first entry)
 XX
 DE p53 serial analysis of gene expression tag #50.
 XX

KW p53: serial analysis of gene expression; SAGE tag; cancer; neoplastic;
 KM rat embryo fibroblast; REF; tumour suppressor; cell cycle control;
 KW tumourigenesis; diagnosis; ss.
 XX
 XX Synthetic.
 OS Rattus sp.
 PN WO9901581-A1.
 XX
 PD 14-JAN-1999.
 XX
 PF 02-JUL-1998: 98WO-US13903.
 XX
 PR 02-JUL-1997: 97US-0051573.
 XX
 PA (GEN2) GENZYME CORP.
 XX
 PI Beaudry GA, Bertelsen AH, Gallella EA, Madden SL;
 XX
 DR WPI: 1999-106079/09.
 XX
 PT Diagnosis of cancer in potentially neoplastic samples - by comparing
 PT the level of transcription between RNA transcripts in two tissue
 PT samples, useful for providing an extensive profile of gene
 PT expression in rat embryo fibroblast (REF) cells
 XX
 PS Example 2; Page 16; 32pp; English.
 XX
 CC A method has been developed for the diagnosis of cancer in potentially
 CC neoplastic samples. The method comprises comparing the level of
 CC transcription between RNA transcripts in two tissue samples (which are
 CC of the same type), where the first sample is potentially neoplastic, and
 CC the second sample is normal human tissue. The first sample is
 CC categorized as neoplastic if its level of transcription is lower than
 CC that of the second sample. The transcript is selected from Alu, RAS,
 CC U6 snRNA, 16S RNA, EGR-1, ribosomal protein S27, ETS-1, 28S RNA, CGR1,
 CC and LIM-2, and it is identified by a tag selected from ribosomal
 CC protein l13a, alpha-tubulin (T1) and (T2), thymosin beta-4, and gamma-
 CC actin. The present sequence represents a serial analysis of gene
 CC expression (SAGE) tag from the present invention. The use of SAGE tags
 CC provides an extensive profile of gene expression in rat embryo
 CC fibroblast (REF) cells containing the (non)-functional p53 tumour
 CC suppression gene. The discovery of new SAGE tags, which are regulated
 CC by p53, enables the diagnosis of genes that are related to cell cycle
 CC control and tumourigenesis.
 CC
 SO Sequence 10 BP; 2 A; 2 C; 3 G; 3 T; 0 other:
 Query Match 100.0%; Score 6; DB 20; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 ID AAX18610/c
 XX AAX18610 standard; DNA; 10 BP.
 AC AAX18610;
 XX
 DT 06-MAY-1999 (first entry)
 XX
 DE p53 serial analysis of gene expression tag #27.
 XX
 DE p53: serial analysis of gene expression; SAGE tag; cancer; neoplastic;
 KM rat embryo fibroblast; REF; tumour suppressor; cell cycle control;
 KW tumourigenesis; diagnosis; ss.
 XX
 OS Synthetic.
 OS Rattus sp.

```

XX MO9901581-A1.
XX 14-JAN-1999.
XX
XX 02-JUL-1998; 98WO-US13903.
XX
XX 02-JUL-1997; 97US-0051573.
XX
XX (GENZ ) GENZYME CORP.
XX
XX Beaudy GA, Bertelsen AH, Galella EA, Madden SL;
XX
XX WPI; 1999-106079/09.
XX
XX Diagnosis of cancer in potentially neoplastic samples - by comparing
XX the level of transcription between RNA transcripts in two tissue
XX samples, useful for providing an extensive profile of gene
XX expression in rat embryo fibroblast (REF) cells
XX
XX Claim 10; Page 15; 32pp; English.
XX
XX A method has been developed for the diagnosis of cancer in potentially
XX neoplastic samples. The method comprises comparing the level of
XX transcription between RNA transcripts in two tissue samples (which are
XX of the same type), where the first sample is potentially neoplastic, and
XX the second sample is normal human tissue. The first sample is
XX categorized as neoplastic if its level of transcription is lower than
XX that of the second sample. The transcript is selected from Alu, RAS,
XX U6 snRNA, 16S RNA, EGR-1, ribosomal protein S27, ETS-1, 28S RNA, CGR11,
XX and LIMK-2, and it is identified by a tag selected from ribosomal
XX protein L13a, alpha-tubulin (T1) and (T2), thymosin beta-4, and gamma-
XX actin. The present sequence represents a serial analysis of gene
XX expression (SAGE) tag from the present invention. The use of SAGE tags
XX provides an extensive profile of gene expression in rat embryo
XX fibroblast (REF) cells containing the (non)-functional p53 tumour
XX suppression gene. The discovery of new SAGE tags, which are regulated
XX by p53, enables the diagnosis of genes that are related to cell cycle
XX control and tumorigenesis.
XX
XX Sequence 10 BP; 2 A; 3 C; 2 G; 3 T; 0 other;
XX
XX Query Match 100.0%; Score 6; DB 20; Length 10;
XX Best Local Similarity 83.3%; Pred. No. 1.7e+05;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 ATGACU 6
XX |11111:
DB 8 ATGACT 3
XX
RESULT 12
ID AAA76009 standard; DNA: 10 BP.
XX
XX AAA76009;
XX
XX 25-JAN-2001 (first entry)
XX
XX DNA sequence analysis oligonucleotide #6.
XX
XX DNA sequence analysis; stacking hybridisation; disease-associated allele;
XX primer; ss.
XX
XX Synthetic.
XX
XX US6090549-A.
XX
XX 18-JUL-2000.
XX
XX 13-MAY-1997; 97US-0855372.
XX
XX 16-JAN-1996; 96US-0587332.
XX

```

```

XX (UYCH-) UNIV CHICAGO.
XX
XX Parinov SV, Barski VE, Dubiley SA, Mirzabekov AD, Kirillov EV;
XX
XX WPI; 2000-523756/47.
XX
XX Determining number of repeat base sequences in a target oligonucleotide
XX for diagnosing a disease, by detecting multiple mutation, utilizing
XX continuous or contiguous stacking hybridization
XX
XX Example 3; Fig 6A; 35pp; English.
XX
XX The present sequence is a synthetic oligonucleotide which was used to
XX demonstrate the methods of the invention. These involve analysing DNA
XX sequences by hybridisation with oligonucleotides associated with
XX polyacrylamide matrices, including continuous and/or contiguous stacking
XX hybridisation. This can be used in sequencing, in diagnostic methods
XX where different alleles are associated with a disease, and in sequencing
XX long DNA fragments containing internal repeats, which enables the
XX identification of unique sequences which may flank such repeats.
XX
XX Sequence 10 BP; 2 A; 1 C; 4 G; 3 T; 0 other;
XX
XX Query Match 100.0%; Score 6; DB 21; Length 10;
XX Best Local Similarity 83.3%; Pred. No. 1.7e+05;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 ATGACU 6
XX |11111:
DB 3 ATGACT 8
XX
RESULT 13
ID AAA76025 standard; DNA: 10 BP.
XX
XX AAA76025;
XX
XX 25-JAN-2001 (first entry)
XX
XX DNA sequence analysis oligonucleotide #16.
XX
XX DNA sequence analysis; stacking hybridisation; disease-associated allele;
XX primer; ss.
XX
XX Synthetic.
XX
XX US6090549-A.
XX
XX 18-JUL-2000.
XX
XX 13-MAY-1997; 97US-0855372.
XX
XX 16-JAN-1996; 96US-0587332.
XX
XX (UYCH-) UNIV CHICAGO.
XX
XX Parinov SV, Barski VE, Dubiley SA, Mirzabekov AD, Kirillov EV;
XX
XX WPI; 2000-523756/47.
XX
XX Determining number of repeat base sequences in a target oligonucleotide
XX for diagnosing a disease, by detecting multiple mutation, utilizing
XX continuous or contiguous stacking hybridization
XX
XX Example 3; Column 12; 35pp; English.
XX
XX The present sequence is a synthetic oligonucleotide which was used to
XX demonstrate the methods of the invention. These involve analysing DNA
XX sequences by hybridisation with oligonucleotides associated with
XX polyacrylamide matrices, including continuous and/or contiguous stacking
XX hybridisation. This can be used in sequencing, in diagnostic methods
XX

```

CC where different alleles are associated with a disease, and in sequencing
 CC long DNA fragments containing internal repeats, which enables the
 CC identification of unique sequences which may flank such repeats.
 XX
 SQ Sequence 10 BP; 2 A; 1 C; 4 G; 3 T; 0 other;

Query Match 100.0%; Score 6; DB 21; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 3 ATGACT 8

RESULT 14
 AA277646
 ID AA277646 standard; DNA; 10 BP.

AC AA277646;

DT 10-APR-2000 (first entry)

DE Human dendritic cell SAGE tag, SEQ ID NO:74.

XX SAGE tag; serial analysis of gene expression; antigen-presenting cell;

KW APC; monocyte-derived dendritic cell; differential gene expression;

KW immunostimulatory cofactor; costimulatory factor; CTL;

KW cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.

XX Homo sapiens.

PN WO965924-A2.

XX 23-DEC-1999.

PF 18-JUN-1999; 99WO-US13800.

XX 19-JUN-1998; 98US-0089833.

PR 19-JUN-1998; 98US-0089844.

PR 19-JUN-1998; 98US-0089853.

PR 19-JUN-1998; 98US-0089878.

PR 19-JUN-1998; 98US-0089991.

PR 19-JUN-1998; 98US-0089992.

PR 19-JUN-1998; 98US-0089993.

PR 19-JUN-1998; 98US-0089994.

PR 19-JUN-1998; 98US-0089997.

PR 19-JUN-1998; 98US-0089999.

PR 19-JUN-1998; 98US-0090000.

PR 19-JUN-1998; 98US-0090035.

PR 19-JUN-1998; 98US-0090036.

PR 19-JUN-1998; 98US-0090039.

PR 19-JUN-1998; 98US-0090040.

PR 19-JUN-1998; 98US-0090041.

PR 19-JUN-1998; 98US-0090042.

PR 19-JUN-1998; 98US-0090043.

PR 19-JUN-1998; 98US-0090044.

PR 19-JUN-1998; 98US-0090045.

PR 19-JUN-1998; 98US-0090047.

PR 19-JUN-1998; 98US-0090048.

PR 19-JUN-1998; 98US-0090072.

PR 19-JUN-1998; 98US-0090076.

PR 19-JUN-1998; 98US-0090077.

PR 19-JUN-1998; 98US-0090078.

PR 19-JUN-1998; 98US-0090079.

PR 19-JUN-1998; 98US-0090080.

PR 08-DEC-1998; 98US-0111715.

XX (GENZ) GENZYME CORP.

PA (ROBE) ROBERTS B L.

PA (SHAN/) SHANKARA S.

XX Roberts BL, Shankara S;

XX
 DR MPI; 2000-106077/09.

PT Isolated polynucleotides differentially expressed in antigen-presenting
 PT cells, useful in gene vaccines against cancer -

XX Claim 1; Page 65; 130pp; English.

CC Sequences AA277573-279709 represent SAGE (serial analysis of gene
 CC expression) tags used to identify mRNA transcripts encoding
 CC immunostimulatory cofactor proteins which are preferentially or
 CC differentially expressed in monocyte-derived dendritic cells compared
 CC with monocytes. Some of the transcripts correspond to known genes or
 CC ESTs (expressed sequence tags) which were previously unknown to be
 CC preferentially or differentially expressed in dendritic cells, while
 CC other transcripts correspond to novel genes. Antigen-presenting cell
 CC (APC)-associated costimulatory factors play an important role in the
 CC activation of the cytotoxic immune response, particularly against tumour
 CC cells. Tumour antigen presentation via the MHC (major histocompatibility
 CC complex) and subsequent recognition by T-cell receptors is alone
 CC insufficient to activate a robust cytotoxic immune response that can
 CC lyse the tumour cells, immunostimulatory cofactors also being required
 CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
 CC sequences identified using the SAGE tags have several potential uses.
 CC They may be used in vaccines to induce an immune response, particularly
 CC against a tumour antigen, to modulate the genotype of an APC, to screen
 CC for agents that modulate expression of differentially expressed genes in
 CC an APC; and as hybridisation probes/amplification primers for the
 CC diagnosis, prognosis and monitoring of diseases related to abnormal
 CC expression of these genes. Detection of the dendritic cell
 CC differentially expressed genes, or of their encoded proteins, can be used
 CC to identify cells as belonging to the monocyte lineage. Cells containing
 CC these genes can be used in active immunotherapy (or to stimulate
 CC production of a population of antigen-specific effector cells) and
 CC vectors containing them are used in gene therapy. Co-administration of
 CC tumour antigens and APC-associated costimulatory factors ensures adequate
 CC antigen presentation to endogenous APCs and upregulates the APCs for the
 CC presentation of co-stimulatory signals, migration to T cell-rich sites,
 CC secretion of T cell growth factors and secretion of chemokines for
 CC recruitment of immune effector cells.

XX Sequence 10 BP; 4 A; 2 C; 2 G; 2 T; 0 other;

SQ Query Match 100.0%; Score 6; DB 21; Length 10;

Best Local Similarity 83.3%; Pred. No. 1.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||:
 Db 5 ATGACT 10

RESULT 15

AA277660/c
 ID AA277660 standard; DNA; 10 BP.

XX AA277660;

DT 10-APR-2000 (first entry)

DE Human dendritic cell SAGE tag, SEQ ID NO:88.

XX SAGE tag; serial analysis of gene expression; antigen-presenting cell;

KW APC; monocyte-derived dendritic cell; differential gene expression;

KW immunostimulatory cofactor; costimulatory factor; CTL;

KW cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.

XX Homo sapiens.

PN WO965924-A2.

XX 23-DEC-1999.

XX

PF 18-JUN-1999; 99WO-US13800.
 XX 19-JUN-1998; 98US-0089833.
 PR 19-JUN-1998; 98US-0089844.
 PR 19-JUN-1998; 98US-0089853.
 PR 19-JUN-1998; 98US-0089878.
 PR 19-JUN-1998; 98US-008991.
 PR 19-JUN-1998; 98US-0089932.
 PR 19-JUN-1998; 98US-0089933.
 PR 19-JUN-1998; 98US-0089934.
 PR 19-JUN-1998; 98US-0089937.
 PR 19-JUN-1998; 98US-0089939.
 PR 19-JUN-1998; 98US-0090000.
 PR 19-JUN-1998; 98US-0090003.
 PR 19-JUN-1998; 98US-0090036.
 PR 19-JUN-1998; 98US-0090039.
 PR 19-JUN-1998; 98US-0090040.
 PR 19-JUN-1998; 98US-0090041.
 PR 19-JUN-1998; 98US-0090042.
 PR 19-JUN-1998; 98US-0090043.
 PR 19-JUN-1998; 98US-0090044.
 PR 19-JUN-1998; 98US-0090045.
 PR 19-JUN-1998; 98US-0090047.
 PR 19-JUN-1998; 98US-0090048.
 PR 19-JUN-1998; 98US-0090072.
 PR 19-JUN-1998; 98US-0090076.
 PR 19-JUN-1998; 98US-0090077.
 PR 19-JUN-1998; 98US-0090078.
 PR 19-JUN-1998; 98US-0090079.
 PR 19-JUN-1998; 98US-0090080.
 PR 08-DEC-1998; 98US-0111715.
 XX (GENZ) GENZYME CORP.
 PA (ROBE/) ROBERTS B L.
 PA (SHAN/) SHANKARA S.
 PI Roberts BL, Shankara S;
 DR WPI: 2000-106077/09.
 XX Isolated polynucleotides differentially expressed in antigen-presenting
 PT cells, useful in gene vaccines against cancer -
 PT
 XX
 PS Claim 1; Page 66; 130pp; English.
 XX Sequences AA27573-279709 represent SAGE (serial analysis of gene
 CC expression) tags used to identify mRNA transcripts encoding
 CC immunostimulatory cofactor proteins which are preferentially or
 CC differentially expressed in monocyte-derived dendritic cells compared
 CC with monocytes. Some of the transcripts correspond to known genes or
 CC ESTs (expressed sequence tags) which were previously unknown to be
 CC preferentially or differentially expressed in dendritic cells, while
 CC other transcripts correspond to novel genes. Antigen-presenting cell
 CC (APC)-associated costimulatory factors play an important role in the
 CC activation of the cytotoxic immune response, particularly against tumour
 CC cells. Tumour antigen presentation via the MHC (major histocompatibility
 CC complex) and subsequent recognition by T-cell receptors is alone
 CC insufficient to activate a robust cytotoxic immune response that can
 CC lyse the tumour cells, immunostimulatory cofactors also being required
 CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
 CC sequences identified using the SAGE tags have several potential uses.
 CC They may be used in vaccines to induce an immune response, particularly
 CC against a tumour antigen; to modulate the genotype of an APC; to screen
 CC for agents that modulate expression of differentially expressed genes in
 CC an APC; and as hybridisation probes/amplification primers for the
 CC diagnosis, prognosis and monitoring of diseases related to abnormal
 CC expression of these genes. Detection of the dendritic cell
 CC differentially expressed genes, or of their encoded proteins, can be used
 CC to identify cells as belonging to the monocyte lineage. Cells containing
 CC these genes can be used in active immunotherapy (or to stimulate
 CC production of a population of antigen-specific effector cells) and
 CC vectors containing them are used in gene therapy. Co-administration of
 CC tumour antigens and APC-associated costimulatory factors ensures adequate

CC antigen presentation to endogenous APCs and upregulates the APCs for the
 CC presentation of co-stimulatory signals, migration to T cell-rich sites,
 CC secretion of T cell growth factors and secretion of chemokines for
 CC recruitment of immune effector cells.

SO Sequence 10 BP; 2 A; 2 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 6; DB 21; Length 10;
 Best Local Similarity 83.3%; Pred. No. 1.7e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
 |||||
 Db 7 ATGACT 2

Search completed: May 21, 2003, 06:54:54
 Job time : 152.25 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:44:53 ; Search time 65 Seconds
(without alignments)
121.889 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6

Sequence: 1 atgacu 6

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 828747 seqs, 660231138 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published_Applications_MA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------------------|---------------------|
| 1 | 6 | 100.0 | 9 | 10 US-09-728-574-14 | Sequence 14, Appl |
| 2 | 6 | 100.0 | 10 | 9 US-10-010-802-278 | Sequence 278, Appl |
| 3 | 6 | 100.0 | 10 | 9 US-09-489-220-18 | Sequence 18, Appl |
| 4 | 6 | 100.0 | 10 | 12 US-10-033-145-74 | Sequence 74, Appl |
| 5 | 6 | 100.0 | 10 | 12 US-10-033-145-88 | Sequence 88, Appl |
| 6 | 6 | 100.0 | 10 | 12 US-10-033-145-98 | Sequence 98, Appl |
| 7 | 6 | 100.0 | 10 | 12 US-10-033-145-795 | Sequence 795, Appl |
| 8 | 6 | 100.0 | 10 | 12 US-10-033-145-1256 | Sequence 1256, Appl |
| 9 | 6 | 100.0 | 10 | 12 US-10-033-145-2043 | Sequence 2043, Appl |
| 10 | 6 | 100.0 | 11 | 9 US-09-249-155-16 | Sequence 16, Appl |
| 11 | 6 | 100.0 | 13 | 9 US-09-818-875-4349 | Sequence 4349, Appl |
| 12 | 6 | 100.0 | 14 | 8 US-08-591-486B-13 | Sequence 13, Appl |
| 13 | 6 | 100.0 | 14 | 8 US-08-591-486B-48 | Sequence 48, Appl |
| 14 | 6 | 100.0 | 14 | 9 US-09-898-027-41 | Sequence 41, Appl |
| 15 | 6 | 100.0 | 15 | 9 US-09-880-887-10 | Sequence 10, Appl |
| 16 | 6 | 100.0 | 15 | 9 US-09-867-915-10 | Sequence 10, Appl |
| 17 | 6 | 100.0 | 15 | 9 US-10-082-476-2 | Sequence 2, Appl |
| 18 | 6 | 100.0 | 15 | 9 US-10-056-414-118 | Sequence 118, Appl |
| 19 | 6 | 100.0 | 15 | 9 US-10-043-875-181 | Sequence 181, Appl |

| | | | | | | | |
|---|----|---|-------|----|----|---------------------|---------------------|
| c | 20 | 6 | 100.0 | 15 | 9 | US-09-848-754A-9112 | Sequence 9112, Appl |
| c | 21 | 6 | 100.0 | 15 | 9 | US-10-010-802-50 | Sequence 50, Appl |
| c | 22 | 6 | 100.0 | 15 | 9 | US-10-010-802-51 | Sequence 51, Appl |
| c | 23 | 6 | 100.0 | 15 | 9 | US-10-010-802-172 | Sequence 172, Appl |
| c | 24 | 6 | 100.0 | 15 | 9 | US-10-010-802-174 | Sequence 174, Appl |
| c | 25 | 6 | 100.0 | 15 | 9 | US-10-010-802-176 | Sequence 176, Appl |
| c | 26 | 6 | 100.0 | 15 | 9 | US-10-010-802-178 | Sequence 178, Appl |
| c | 27 | 6 | 100.0 | 15 | 9 | US-09-439-439-19 | Sequence 19, Appl |
| c | 28 | 6 | 100.0 | 15 | 9 | US-10-287-919-2017 | Sequence 2017, Appl |
| c | 29 | 6 | 100.0 | 15 | 10 | US-09-504-231A-481 | Sequence 481, Appl |
| c | 30 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1100 | Sequence 1100, Appl |
| c | 31 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1220 | Sequence 1220, Appl |
| c | 32 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1221 | Sequence 1221, Appl |
| c | 33 | 6 | 100.0 | 15 | 10 | US-09-504-231A-1242 | Sequence 1242, Appl |
| c | 34 | 6 | 100.0 | 15 | 10 | US-09-179-536B-295 | Sequence 295, Appl |
| c | 35 | 6 | 100.0 | 15 | 10 | US-09-780-954A-9 | Sequence 9, Appl |
| c | 36 | 6 | 100.0 | 15 | 10 | US-09-780-954A-10 | Sequence 10, Appl |
| c | 37 | 6 | 100.0 | 15 | 10 | US-09-274-553D-481 | Sequence 481, Appl |
| c | 38 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1100 | Sequence 1100, Appl |
| c | 39 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1220 | Sequence 1220, Appl |
| c | 40 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1221 | Sequence 1221, Appl |
| c | 41 | 6 | 100.0 | 15 | 10 | US-09-274-553D-1242 | Sequence 1242, Appl |
| c | 42 | 6 | 100.0 | 15 | 10 | US-09-953-242-4 | Sequence 4, Appl |
| c | 43 | 6 | 100.0 | 15 | 10 | US-09-953-242-5 | Sequence 5, Appl |
| c | 44 | 6 | 100.0 | 15 | 10 | US-09-953-242-13 | Sequence 13, Appl |
| c | 45 | 6 | 100.0 | 16 | 8 | US-08-591-486B-11 | Sequence 11, Appl |

ALIGNMENTS

RESULT 1
US-09-728-574-14
; Sequence 14, Application US/09728574
; Patent No. US20020137036A1
; GENERAL INFORMATION:
; APPLICANT: Strategene
; TITLE OF INVENTION: Methods for Detection of a Target Nucleic Acid By Capture
; FILE REFERENCE: 25436/1660
; CURRENT APPLICATION NUMBER: US/09/728, 574
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 09/728574
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Saccharomyces sp.
; FEATURE:
; NAME/KEY: GCN4 DNA binding site
; LOCATION: (1)..(9)
US-09-728-574-14

Query Match
Best local Similarity 100.0%; Score 6; DB 10; Length 9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 1 ATGACT 6

RESULT 2
US-10-010-802-278/c
; Sequence 278, Application US/10010802
; Publication No. US20030078220A1
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals
; APPLICANT: Chew, Anne
; APPLICANT: Denton, R. Rex
; APPLICANT: Duda, Amy
; APPLICANT: Nandabalan, Krishnan

APPLICANT: Stephens, J. Claborn
APPLICANT: Windemuth, Andreas
TITLE OF INVENTION: Drug Target Isogenes: Polymorphisms in the Interleukin
FILE OF INVENTION: 4 Receptor Alpha Gene
FILE REFERENCE: MGH-0002052 IL4R alpha
CURRENT APPLICATION NUMBER: US/10/010,802
CURRENT FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: PCT/US00/19094
PRIOR FILING DATE: 2000-07-13
NUMBER OF SEQ ID NOS: 413
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 278
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-010-802-278

Query Match 100.0%; Score 6; DB 9; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 9 ATGACT 4

RESULT 3
US-09-489-220-18
Sequence 18, Application US/09489220
Patent No. US20020110808A1
GENERAL INFORMATION:
APPLICANT: Reidhaer-Olson, John F.
TITLE OF INVENTION: Toxicant-Induced Differential Gene Expression
FILE REFERENCE: 16528A-0389000S
CURRENT APPLICATION NUMBER: US/09/489,220
CURRENT FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 18
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ten nucleotides
OTHER INFORMATION: following M13R(-48) primer sequence in 5'
OTHER INFORMATION: arbitrary primer (Arp) 14
US-09-489-220-18

Query Match 100.0%; Score 6; DB 10; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 4 ATGACT 9

RESULT 4
US-10-033-145-74
Sequence 74, Application US/10033145
Patent No. US2002015151A1
GENERAL INFORMATION:
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GA0201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137

SOFTWARE: Patentin version 3.0
SEQ ID NO 74
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-74

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 5 ATGACT 10

RESULT 5
US-10-033-145-88/c
Sequence 88, Application US/10033145
Patent No. US2002015151A1
GENERAL INFORMATION:
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GA0201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137
SOFTWARE: Patentin version 3.0
SEQ ID NO 88
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-88

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||
DB 7 ATGACT 2

RESULT 6
US-10-033-145-98
Sequence 98, Application US/10033145
Patent No. US2002015151A1
GENERAL INFORMATION:
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GA0201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137
SOFTWARE: Patentin version 3.0
SEQ ID NO 98
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-98

Query Match 100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;


```
OY      1 ATGACU 6
      |||||:
Db      2 ATGACT 7

RESULT 7
US-10-033-145-795/c
; Sequence 795, Application US/10033145
; Patent No. US2002015151A1
; GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION
; APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKARA, SRINIVAS
; TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
; FILE REFERENCE: GA0201C
; CURRENT APPLICATION NUMBER: US/10/033,145
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/13800
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 2137
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 795
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-145-795

Query Match      100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 ATGACU 6
      |||||:
Db      8 ATGACT 3

RESULT 8
US-10-033-145-1256
; Sequence 1256, Application US/10033145
; Patent No. US2002015151A1
; GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION
; APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKARA, SRINIVAS
; TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
; FILE REFERENCE: GA0201C
; CURRENT APPLICATION NUMBER: US/10/033,145
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/13800
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 2137
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1256
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-145-1256

Query Match      100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 ATGACU 6
      |||||:
Db      2 ATGACT 7

RESULT 9
US-10-033-145-2043
; Sequence 2043, Application US/10033145
; Patent No. US2002015151A1
; GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION
```

```
; APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKARA, SRINIVAS
; TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
; FILE REFERENCE: GA0201C
; CURRENT APPLICATION NUMBER: US/10/033,145
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/13800
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 2137
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2043
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-145-2043

Query Match      100.0%; Score 6; DB 12; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 ATGACU 6
      |||||:
Db      1 ATGACT 6

RESULT 10
US-09-249-155-16
; Sequence 16, Application US/09249155
; Publication No. US20030037345A1
; GENERAL INFORMATION:
; APPLICANT: Heber-Katz, Ellen
; TITLE OF INVENTION: Compositions and Methods for Wound
; FILE REFERENCE: 00486.78503
; CURRENT APPLICATION NUMBER: US/09/249,155
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: 60/074,737
; EARLIER FILING DATE: 1998-02-13
; EARLIER APPLICATION NUMBER: 60/097,937
; EARLIER FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/102,051
; EARLIER FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 16
; LENGTH: 11
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-249-155-16

Query Match      100.0%; Score 6; DB 9; Length 11;
Best Local Similarity 83.3%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 ATGACU 6
      |||||:
Db      1 ATGACT 6

RESULT 11
US-09-818-875-4349
; Sequence 4349, Application US/09818875
; Publication No. US20030051270A1
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric B.
; APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
; FILE REFERENCE: Napro-4
; CURRENT APPLICATION NUMBER: US/09/818,875
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/192,176
```

PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/192,179
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: US 60/208,538
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/244,989
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 4385
SOFTWARE: Friedman macro Napro4
SEQ ID NO 4349
LENGTH: 13
TYPE: DNA
ORGANISM: Escherichia coli
US-09-818-875-4349

Query Match
Best Local Similarity 83.3%; Pred. No. 1.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 8 ATGACT 13

RESULT 12
US-08-591-486B-13/C
Sequence 13, Application US/08591486B
Patent No. US20020037866A1
GENERAL INFORMATION:
APPLICANT: Schlingensiepen, Georg F
APPLICANT: Schlingensiepen, Reimar
APPLICANT: Schlingensiepen, Karl-Hermann
APPLICANT: Göttingen, Wolfgang Brysch
TITLE OF INVENTION: A Pharmaceutical Composition
TITLE OF INVENTION: Compising Antisense-Nucleic Acid for Prevention and/or Treat
TITLE OF INVENTION: of Neuronal Injury, Degeneration and Cell Death and for the
TITLE OF INVENTION: Treatment of Neoplasms
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern
STREET: 400 Seventh Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,486B
FILING DATE: 11-JAN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93111059.7
FILING DATE: 10-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02218
FILING DATE: 6-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: 10496/P60122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-9350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: YES
US-08-591-486B-13

Query Match
Best Local Similarity 100.0%; Score 6; DB 8; Length 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 9 ATGACT 4

RESULT 13
US-08-591-486B-48
Sequence 48, Application US/08591486B
Patent No. US20020037866A1
GENERAL INFORMATION:
APPLICANT: Schlingensiepen, Georg F
APPLICANT: Schlingensiepen, Reimar
APPLICANT: Schlingensiepen, Karl-Hermann
APPLICANT: Göttingen, Wolfgang Brysch
TITLE OF INVENTION: A Pharmaceutical Composition
TITLE OF INVENTION: Compising Antisense-Nucleic Acid for Prevention and/or Tre
TITLE OF INVENTION: of Neuronal Injury, Degeneration and Cell Death and for the
TITLE OF INVENTION: Treatment of Neoplasms
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern
STREET: 400 Seventh Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,486B
FILING DATE: 11-JAN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93111059.7
FILING DATE: 10-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02218
FILING DATE: 6-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: 10496/P60122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-9350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: YES
US-08-591-486B-48

Query Match
Best Local Similarity 100.0%; Score 6; DB 8; Length 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 9 ATGACT 14

RESULT 14

US-09-998-027-41/C
; Sequence 41, Application US/09998027
; Publication No. US20030093819A1
; GENERAL INFORMATION:
; APPLICANT: D'Andrea et al.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Diagnosis and treatment of Cancers Associated with Defective
; TITLE OF INVENTION: DNA Repair Mechanisms
; FILE REFERENCE: 2486/101
; CURRENT APPLICATION NUMBER: US/09/998,027
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; OTHER INFORMATION: Intron/Exon Junctions of FANCD
US-09-998-027-41

Query Match

100.0%; Score 6; DB 9; Length 14;

Best Local Similarity 83.3%; Pred. No. 1.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 12 ATGACT 7

RESULT 15

US-09-880-887-10/C
; Sequence 10, Application US/09880887
; Patent No. US20020165177A1
; GENERAL INFORMATION:
; APPLICANT: NEGRIER, CLAUDE
; APPLICANT: PLANTIER, JEAN LUC
; TITLE OF INVENTION: MODIFIED FACTOR VIII CDNA
; FILE REFERENCE: 06478.1441
; CURRENT APPLICATION NUMBER: US/09/880,887
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 09/526,935
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: EP 99104050.2
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
US-09-880-887-10

Query Match

100.0%; Score 6; DB 9; Length 15;

Best Local Similarity 83.3%; Pred. No. 1.6e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGACU 6
|||||:
Db 8 ATGACT 3

Search completed: May 21, 2003, 08:44:51
Job time : 65 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 06:28:43 ; Search time 34.75 Seconds
(without alignments)
52.951 Million cell updates/sec

Title: US-09-936-146-1

Perfect score: 6

Sequence: 1 atgacg 6

Scoring table: OLIGO_NUC

Searched: 441362 seqs, 153338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1 | 6 | 100.0 | 8 | US-08-859-954-66 | Sequence 66, Appl |
| 2 | 6 | 100.0 | 3 | US-08-859-954-158 | Sequence 158, App |
| 3 | 6 | 100.0 | 10 | US-08-388-353-232 | Sequence 232, App |
| 4 | 6 | 100.0 | 10 | US-08-388-353-233 | Sequence 233, App |
| 5 | 6 | 100.0 | 10 | US-08-388-353-234 | Sequence 234, App |
| 6 | 6 | 100.0 | 10 | US-08-388-353-235 | Sequence 235, App |
| 7 | 6 | 100.0 | 10 | US-08-388-353-236 | Sequence 236, App |
| 8 | 6 | 100.0 | 10 | US-08-488-551B-232 | Sequence 232, App |
| 9 | 6 | 100.0 | 10 | US-08-488-551B-233 | Sequence 233, App |
| 10 | 6 | 100.0 | 10 | US-08-488-551B-234 | Sequence 234, App |
| 11 | 6 | 100.0 | 10 | US-08-488-551B-235 | Sequence 235, App |
| 12 | 6 | 100.0 | 10 | US-08-488-551B-236 | Sequence 236, App |
| 13 | 6 | 100.0 | 10 | US-08-859-954-158 | Sequence 158, App |
| 14 | 6 | 100.0 | 10 | US-08-859-954-159 | Sequence 159, App |
| 15 | 6 | 100.0 | 10 | US-08-859-954-160 | Sequence 160, App |
| 16 | 6 | 100.0 | 10 | US-08-859-954-161 | Sequence 161, App |
| 17 | 6 | 100.0 | 10 | US-08-859-954-162 | Sequence 162, App |
| 18 | 6 | 100.0 | 10 | US-08-859-954-163 | Sequence 163, App |
| 19 | 6 | 100.0 | 10 | US-08-859-954-164 | Sequence 164, App |
| 20 | 6 | 100.0 | 10 | US-08-859-954-165 | Sequence 165, App |
| 21 | 6 | 100.0 | 10 | US-08-859-954-166 | Sequence 166, App |
| 22 | 6 | 100.0 | 10 | US-08-859-954-167 | Sequence 167, App |
| 23 | 6 | 100.0 | 10 | US-08-859-954-168 | Sequence 168, App |
| 24 | 6 | 100.0 | 10 | US-08-859-954-169 | Sequence 169, App |
| 25 | 6 | 100.0 | 10 | US-08-859-954-170 | Sequence 170, App |
| 26 | 6 | 100.0 | 10 | US-08-859-954-171 | Sequence 171, App |
| 27 | 6 | 100.0 | 10 | US-08-859-954-172 | Sequence 172, App |

| | | | | | | |
|------|---|-------|----|---|-------------------|--------------------|
| C 28 | 6 | 100.0 | 12 | 4 | US-09-631-349A-12 | Sequence 12, Appl |
| C 29 | 6 | 100.0 | 13 | 2 | US-08-867-820A-49 | Sequence 49, Appl |
| C 30 | 6 | 100.0 | 13 | 4 | US-08-797-812-16 | Sequence 16, Appl |
| C 31 | 6 | 100.0 | 13 | 6 | 5225337-13 | Patent No. 5225337 |
| C 32 | 6 | 100.0 | 14 | 1 | US-08-303-004-1 | Sequence 1, Appl |
| C 33 | 6 | 100.0 | 14 | 1 | US-08-358-810A-3 | Sequence 3, Appl |
| C 34 | 6 | 100.0 | 14 | 1 | US-08-484-712A-3 | Sequence 3, Appl |
| C 35 | 6 | 100.0 | 14 | 1 | US-08-359-295C-16 | Sequence 16, Appl |
| C 36 | 6 | 100.0 | 14 | 2 | US-08-485-105A-16 | Sequence 16, Appl |
| C 37 | 6 | 100.0 | 14 | 2 | US-08-769-945C-2 | Sequence 2, Appl |
| C 38 | 6 | 100.0 | 14 | 2 | US-08-769-945C-3 | Sequence 3, Appl |
| C 39 | 6 | 100.0 | 14 | 2 | US-08-769-945C-4 | Sequence 4, Appl |
| C 40 | 6 | 100.0 | 14 | 2 | US-08-769-945C-15 | Sequence 15, Appl |
| C 41 | 6 | 100.0 | 14 | 2 | US-08-769-945C-17 | Sequence 17, Appl |
| C 42 | 6 | 100.0 | 14 | 2 | US-08-769-945C-18 | Sequence 18, Appl |
| C 43 | 6 | 100.0 | 14 | 2 | US-08-769-945C-19 | Sequence 19, Appl |
| C 44 | 6 | 100.0 | 14 | 3 | US-09-183-650-16 | Sequence 16, Appl |
| C 45 | 6 | 100.0 | 14 | 4 | US-09-081-646-7 | Sequence 7, Appl |

ALIGNMENTS

RESULT 1
US-08-859-954-66/c
Sequence 66, Application US/08859954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramin
TITLE OF INVENTION: Design and Optimized Primer Library for
TITLE OF INVENTION: Gene Sequencing and Method Thereof
NUMBER OF SEQUENCES: 566
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,954
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REFERENCE/DOCKET NUMBER: D-5900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5246
TELEFAX: 713/651-5246
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligonucleotide"
HYPOTHETICAL: YES
ANTI-SENSE: YES
US-08-859-954-66
Query Match 100.0%, Score 6, DB 3, Length 8;

Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 8 ATGACT 3

RESULT 2

US-08-959-954-158
Sequence 158, Application US/08859954
Patent No. 6083695
GENERAL INFORMATION:
APPLICANT: Hardin, Susan H.
APPLICANT: Homayouni, Ramtin
APPLICANT: Hardin, Paul E.
TITLE OF INVENTION: Design and Optimized Primer Library for
TITLE OF INVENTION: Gene Sequencing and Method Thereof
NUMBER OF SEQUENCES: 566
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,954
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligonucleotide"
HYPOTHETICAL: YES
ANTI-SENSE: YES
US-08-959-954-158

Query Match 100.0%; Score 6; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 3.6e+07;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 2 ATGACT 7

RESULT 3

US-08-388-353-232
Sequence 232, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.

APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 232:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-232

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
DB 5 ATGACT 10

RESULT 4

US-08-388-353-233
Sequence 233, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353

FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 233:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-233

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 4 ATGACT 9

RESULT 5
US-08-388-353-234
Sequence 234, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 234:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-234

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 3 ATGACT 8

RESULT 6
US-08-388-353-235
Sequence 235, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 235:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-235

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 2 ATGACT 7

RESULT 7
US-08-388-353-236
Sequence 236, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David

TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 garden city plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: DIGIGLIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-236

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGACU 6
Db 1 ATGACT 6

RESULT 8
US-08-488-551B-232
Sequence 232, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)

FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 232:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-232

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGACU 6
Db 5 ATGACT 10

RESULT 9
US-08-488-551B-233
Sequence 233, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366

INFORMATION FOR SEQ ID NO: 233:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-233

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGACU 6
Db 4 ATGACT 9

RESULT 10
US-08-488-551B-234
Sequence 234, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488, 551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 234:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-234

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGACU 6
Db 3 ATGACT 8

RESULT 11
US-08-488-551B-235
Sequence 235, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488, 551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 235:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-235

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGACU 6
Db 2 ATGACT 7

RESULT 12
US-08-488-551B-236
Sequence 236, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
APPLICANT: David Cooper

TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-236

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
11111:
Db 1 ATGACT 6

RESULT 13
US-08-855-372B-19
Sequence 19, Application US/08855372B
Patent No. 6090549
GENERAL INFORMATION:
APPLICANT: Mirzabekov, Andrei D
APPLICANT: Parinov, Sergei V
APPLICANT: Barsky, Victor E
APPLICANT: Kirillov, Eugene V
APPLICANT: Dubilay, Svetlana A
TITLE OF INVENTION: Use of Continuous/Contiguous Stacking Hybridization as a Diagn
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHERSKOV & FLAYNIK
STREET: 20 N. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch, 1.4 MB storage

COMPUTER: PC
OPERATING SYSTEM: Microsoft Windows 98
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,372B
FILING DATE: 13-MAY-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/587,332
FILING DATE: 16-JAN-96
ATTORNEY/AGENT INFORMATION:
NAME: Cherskov, Michael J
REGISTRATION NUMBER: 33,664
REFERENCE/DOCKET NUMBER: ANL-IN-95-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 621-1330
TELEFAX: (312) 621-0088
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 bases
TYPE: nucleic acid
STRANDEDNESS: No. 6090549 Applicable
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: yes
US-08-855-372B-19

Query Match 100.0%; Score 6; DB 3; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
11111:
Db 3 ATGACT 8

RESULT 14
US-08-855-372B-87
Sequence 87, Application US/08855372B
Patent No. 6090549
GENERAL INFORMATION:
APPLICANT: Mirzabekov, Andrei D
APPLICANT: Parinov, Sergei V
APPLICANT: Barsky, Victor E
APPLICANT: Kirillov, Eugene V
APPLICANT: Dubilay, Svetlana A
TITLE OF INVENTION: Use of Continuous/Contiguous Stacking Hybridization as a Di
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHERSKOV & FLAYNIK
STREET: 20 N. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch, 1.4 MB storage
COMPUTER: PC
OPERATING SYSTEM: Microsoft Windows 98
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,372B
FILING DATE: 13-MAY-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/587,332
FILING DATE: 16-JAN-96
ATTORNEY/AGENT INFORMATION:
NAME: Cherskov, Michael J
REGISTRATION NUMBER: 33,664
REFERENCE/DOCKET NUMBER: ANL-IN-95-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 621-1330
TELEFAX: (312) 621-0088
INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 bases
TYPE: nucleic acid
STRANDEDNESS: No. 6090549 Applicable
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: yes
US-08-855-372B-87

Query Match
Best Local Similarity 83.3%; Score 6; DB 3; Length 10;
Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 3 ATGACT 8

RESULT 15
US-08-991-525B-87
Sequence 87, Application US/08991525B
Patent No. 6093811
GENERAL INFORMATION:
APPLICANT: Bennett and Mirabelli
TITLE OF INVENTION: Oligonucleotide Modulation
TITLE OF INVENTION: of Cell Adhesion
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: Windows 95
SOFTWARE: WORDPERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,525B
FILING DATE: December 16, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 440,740
FILING DATE: May 12, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 063,167
FILING DATE: May 17, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 969,151
FILING DATE: February 10, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 007,997
FILING DATE: January 21, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 939,855
FILING DATE: September 2, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 567,286
FILING DATE: August 14, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515
TELEFAX: (856) 810-1454
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 10
TYPE: Nucleic Acid
STRANDEDNESS: Single

TOPOLOGY: linear
ANTI-SENSE: yes
US-08-991-525B-87

Query Match
Best Local Similarity 83.3%; Score 6; DB 3; Length 10;
Pred. No. 2.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACU 6
|||||
Db 5 ATGACT 10

Search completed: May 21, 2003, 08:40:17
Job time : 34.75 secs

FEATURES
Source
Location/Qualifiers
1. .15
/organism="Solanum tuberosum"
/cultivar="Kennebec or Binje"
/db.xref="taxon:4113"
/clone="STM018"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
, tubers, or roots."

BASE COUNT 3 a 5 c 2 g 5 t
ORIGIN

Query Match 100.0%; Score 6; DB 14; Length 15;
Best Local Similarity 83.3%; Pred. No. 5.5e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCU 6
Db 9 GCCTCT 14

RESULT 2
BQ789979 16 bp mRNA linear EST 30-JUL-2002
LOCUS
DEFINITION BQ789979 Heterobasidion annosum - Scots pine infection stage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage005a12, mRNA sequence.
BQ789979.1 GI:22004941
EST.
Pinus sylvestris/Heterobasidion annosum.
Pinus sylvestris/Heterobasidion annosum
Eukaryota; mixed EST libraries.
1 (bases 1 to 16)
Asiegbu, F.O., Nahalkova, J. and Dean, R.A.
Selected Expressed sequence tags of cDNA clones from the
interaction of the root rot fungus (Heterobasidion annosum) with
seedling roots of Scots pine (Pinus sylvestris)
Unpublished (2001)
Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala,
Sweden
Tel: +46 18 67 15 98
Fax: +46 18 30 92 45
Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer.
Location/Qualifiers
1. .16
/organism="Pinus sylvestris/Heterobasidion annosum"
/db.xref="taxon:169015"
/clone="hage005a12"
/clone_lib="Heterobasidion annosum - Scots pine infection
stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for
6 days with H. annosum"
/note="Vector: pT-Adv; Site_1: EcoRI; The subtractive
hybridization cDNA library was constructed from scots pine
roots infected for 6-days with mycelia of Heterobasidion
annosum (FP5)."

BASE COUNT 5 a 3 c 3 g 5 t
ORIGIN

Query Match 100.0%; Score 6; DB 14; Length 16;
Best Local Similarity 83.3%; Pred. No. 5.7e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCU 6
Db 7 GCCTCT 12

RESULT 3
A1569191 19 bp mRNA linear EST 14-MAY-1999
LOCUS
DEFINITION A1569191 NCI-CGAP Panel Homo sapiens cDNA clone IMAGE:224943 3'
tr83f12.x1 NCI-CGAP Panel Homo sapiens cDNA clone IMAGE:224943 3'
similar to SM:PCCE_HUMAN P28074 PROTEASOME EPSILON CHAIN PRECURSOR
; mRNA sequence.
A1569191.1 GI:4532565
EST.
A1569191.1 GI:4532565
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 19)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Trace considered overall poor quality
Insert Length: 1117 Std Error: 0.00
Seq primer: -40UP from G1bco
High quality sequence stop: 1
POLYA-NO.

FEATURES
Source
Location/Qualifiers
1. .19
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:224943"
/clone_lib="NCI-CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH108"
/note="Organ: pancreas; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 2 a 7 c 5 g 5 t
ORIGIN

Query Match 100.0%; Score 6; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCTCU 6
Db 8 GCCTCT 13

RESULT 4
A2429998 19 bp DNA linear GSS 03-OCT-2000
LOCUS
DEFINITION A2429998 Mouse 10kb plasmid UGCM1 library Mus musculus genomic
clone UGCM10214F16 F, DNA sequence.
A2429998
A2429998.1 GI:10554011
GSS.
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0214 row: F column: 16
Seq primer: CGTTGTAAACGACGCCACT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0214F16"
/clone_1lb="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|pb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
5 a 3 c 9 g 2 t

ORIGIN

Query Match
100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:

Db 11 GCCTCT 6

RESULT 5
A2475079 19 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION
IM0293B17F Mouse 10kb plasmid UGGCM library Mus musculus genomic
clone UGGCM0293B17 F, DNA sequence.
ACCESSION
A2475079
VERSION
A2475079.1 GI:10633204
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0293 row: B column: 17
Seq primer: CGTTGTAAACGACGCCACT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0293B17"
/clone_1lb="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|pb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 4 c 9 g 0 t

ORIGIN

Query Match
100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:

Db 18 GCCTCT 13

RESULT 6
A2480905 19 bp DNA linear GSS 04-OCT-2000
LOCUS
DEFINITION
IM0302N22R Mouse 10kb plasmid UGGCM library Mus musculus genomic
clone UGGCM0302N22 R, DNA sequence.
ACCESSION
A2480905
VERSION
A2480905.1 GI:10642066
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
'M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Bldg. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 22
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0302N22"
/clone_1lb="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321141gb1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
1 a 5 c 9 g 4 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 7 GCCTCT 12

RESULT 7
A2509071/c 19 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0351A21R Mouse 10kb plasmid UGC1M library Mus musculus genomic
DEFINITION clone UGC1M0351A21 R, DNA sequence.
ACCESSION A2509071 GI:10690387
VERSION A2509071.1
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
'M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Bldg. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0351 row: A column: 21
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. .19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0351A21"
/clone_1lb="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321141gb1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
4 a 4 c 9 g 2 t

ORIGIN
Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||
Db 13 GCCTCT 8

RESULT 8
A2651803 19 bp DNA linear GSS 14-DEC-2000
LOCUS 1M0522N11R Mouse 10kb plasmid UGC1M library Mus musculus genomic
DEFINITION clone UGC1M0522N11 R, DNA sequence.
ACCESSION A2651803 GI:11787672
VERSION A2651803.1
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
.M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0522 row: N column: 11
Seq primer: CACACAGGAACAGCATATGACC
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0537N06"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 7 c 5 g 4 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTCU 6
|||||

Db 7 GCCCTCT 12

RESULT 9
AZ659603/c 19 bp DNA 1linear GSS 14-DEC-2000
LOCUS
DEFINITION 1M0537N06F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
clone UUCG1M0537N06 F, DNA sequence.
ACCESSION
VERSION AZ659603
KEYWORDS AZ659603.1 GI:11796749
SOURCE GSS.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
.M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0537 row: N column: 06
Seq primer: CGTGTAAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 19.

FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0537N06"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
7 a 2 c 7 g 3 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTCU 6
|||||

Db 12 GCCCTCT 7

RESULT 10
AZ800646/c 19 bp DNA 1linear GSS 16-FEB-2001
LOCUS
DEFINITION 2M0058H14R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
clone UUCG2M0058H14 R, DNA sequence.
ACCESSION
VERSION AZ800646
KEYWORDS AZ800646.1 GI:12952969
SOURCE GSS.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0058 row: H column: 14
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0058H14"
/clone_1lb="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114(gb)AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
3 a 4 c 7 g 5 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 7 GCCTCT 2

RESULT 11
A2834038 19 bp DNA linear GSS 20-FEB-2001
LOCUS 2M0116H01R Mouse 10kb plasmid U06C1M library Mus musculus genomic
DEFINITION clone U06C2M0116H01 R, DNA sequence.
ACCESSION A2834038
VERSION A2834038.1 GI:13003946
KEYWORDS GSS.
SOURCE house mouse;
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0116 row: H column: 01
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0116H01"
/clone_1lb="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114(gb)AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
6 a 2 c 7 g 4 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 19;
Best Local Similarity 83.3%; Pred. No. 6.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTCU 6
|||||:
Db 6 GCCTCT 1

RESULT 12
B0789776 20 bp mRNA linear EST 30-JUL-2002
LOCUS B0789776
DEFINITION hage002aa10 Heterobasidion annosum - Scots pine infection atage
(HAGE) subtraction cDNA library Pinus sylvestris/Heterobasidion
annosum cDNA clone hage002aa10, mRNA sequence.
ACCESSION B0789776
VERSION B0789776.1 GI:22004738
KEYWORDS EST.
SOURCE Pinus sylvestris/Heterobasidion annosum.
ORGANISM Pinus sylvestris/Heterobasidion annosum
REFERENCE Eukaryota; mixed EST libraries.
1 (bases 1 to 20)

AUTHORS Asiegbu, F.O., Nahalkova, J. and Dean, R.A.

TITLE

Selected Expressed sequence tags of cDNA clones from the interaction of the root rot fungus (*Heterobasidion annosum*) with seedling roots of Scots pine (*Pinus sylvestris*)

JOURNAL

Unpublished (2001)

COMMENT

Contact: Fred O. Asiegbu
Dept. of Forest Mycology & Pathology
Swedish University of Agriculture, Box 7026, S-750 07, Uppsala, Sweden

Tel: +46 18 67 15 98
Fax: +46 18 30 92 45

Email: Fred.Asiegbu@mykopat.slu.se
Seq primer: T7 primer.

FEATURES

Location/Qualifiers

1..20
/organism="Pinus sylvestris/Heterobasidion annosum"
/db_xref="taxon:169015"
/clone="hage002a10"
/clone_1id="Heterobasidion annosum - Scots pine infection stage (HAGE) subtraction cDNA library"
/dev_stage="Seedling roots of scots pine were infected for 6 days with H. annosum"
/note="Vector: pT-Adv; Site_1: EcoRI; The subtractive hybridization cDNA library was constructed from scots pine roots infected for 6-days with mycelia of *Heterobasidion annosum* (FP5)."
BASE COUNT 4 a 5 c 3 g 7 t 1 others
ORIGIN

Query Match 100.0%; Score 6; DB 14; Length 20;

Best Local Similarity 83.3%; Pred. No. 6.4e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCT 6
|||||
DB 12 GCCTCT 17

RESULT 13

LOCUS

AZ309156 20 bp DNA linear GSS 29-SEP-2000

DEFINITION clone UDC1M0013B09 F, DNA sequence.

ACCESSION AZ309156

VERSION AZ309156.1 GI:10349862

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0013 row: B column: 09

Seq primer: CCGTGTAAACGACGCCACT
Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers
1..20

FEATURES

Source

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UDC1M0013B09"

/clone_1id="Mouse 10kb plasmid UDC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adapter oligonucleotides were ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1) a copy-number

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 11 c 4 g 3 t
ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 20;

Best Local Similarity 83.3%; Pred. No. 6.4e+05;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCT 6
|||||
DB 3 GCCTCT 8

RESULT 14

LOCUS

AZ366535 20 bp DNA linear GSS 02-OCT-2000

DEFINITION IM0115M1SR Mouse 10kb plasmid UDC1M library Mus musculus genomic

clone UDC1M0115M1S R, DNA sequence.

ACCESSION AZ366535

VERSION AZ366535.1 GI:10480235

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0115 row: M column: 15

Seq primer: CACACAGAAACACCTATGACC
Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers
1..20

FEATURES

Source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0115M15"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

BASE COUNT 5 a 6 c 3 g 6 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 20;
 Best Local Similarity 83.3%; Pred. NO. 6.4e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6

DB 14 GCCTCT 19

RESULT 15

AZ615164

LOCUS 20 bp DNA linear GSS 13-DEC-2000

DEFINITION 1M0444B24F Mouse 10kb plasmid U06C1M library Mus musculus genomic

ACCESSION AZ615164

VERSION AZ615164.1 GI:11737354

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 20) Dunn, P., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,

'M., Rose, M., Rose, R., Stokes, R., Tingley, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

CONTACT: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: dunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0444 row: B column: 24

Seq primer: CGTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers

1. 20

FEATURES

source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0444B24"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

BASE COUNT 3 a 7 c 5 g 5 t

ORIGIN

Query Match 100.0%; Score 6; DB 17; Length 20;
 Best Local Similarity 83.3%; Pred. NO. 6.4e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTCU 6

DB 7 GCCTCT 12

Search completed: May 21, 2003, 06:26:02

Job time : 1097 secs